THE DISSERTATION COMMITTEE FOR DOROTHY LAM MORGAN CERTIFIES THAT THIS IS THE APPROVED VERSION OF THE FOLLOWING DISSERTATION:

A SPATIAL ECONOMETRIC APPROACH TO THE STUDY OF SOCIAL INFLUENCE

Committee:

Tse-min Lin, Supervisor

Robert Luskin

Brian Roberts

Daron Shaw

Samuel Workman

A SPATIAL ECONOMETRIC APPROACH TO THE STUDY OF SOCIAL INFLUENCE

BY

DOROTHY LAM MORGAN, B.A.; M.A.; M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

DECEMBER 2012

DEDICATION

For my loving husband Clint.

ACKNOWLEDGEMENTS

I wish to thank various people for their contribution to this project. First, many thanks to Tse-min Lin, my dissertation supervisor. He inspired many of the ideas for this dissertation; in particular, Chapter 3 is based on one of his ongoing research projects and should be considered a collaborative chapter. The completion of this dissertation would not have been possible without his invaluable guidance and feedback.

I also want to thank Robert Luskin, for his encouragement and advice; the rest of my committee members for their comments and suggestions; and Annette Carlile for her resourcefulness and patience.

I also wish to thank Clint Morgan, Pete Mohanty, and Mary Selph for their love, friendship, and support.

A SPATIAL ECONOMETRIC APPROACH TO THE STUDY OF SOCIAL INFLUENCE

Dorothy Lam Morgan, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Tse-min Lin

While political scientists have traditionally examined social influence through social network or contextual studies, this dissertation argues for the use of spatial econometrics as an alternative approach. While spatial econometrics is not new to political science, the dissertation attempts to broaden its application by exploring spaces based on geography, demographic characteristics, and ideology. Social influence can be understood as a form of spatial interdependence among individuals in these spaces and can be analyzed as spatial autocorrelation.

In the dissertation, I discuss the dimensions of the three spaces, what might account for mutual influence in these spaces, how to measure distances in these spaces, and how to use these distances for estimating social influence in models of political attitudes using ANES data. By taking a broader approach to space, I show that spatial econometrics can offer many advantages over more conventional approaches.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Announcing that the "GOP is the new black" and that "Martin Luther King, Jr. was a Republican," three billboards appeared in Austin, Texas in October 2011. Conservative group Raging Elephants had put up these billboards to encourage blacks to think about voting Republican and to weaken what it saw as an undue loyalty to the Democratic Party. According to Raging Elephant's representatives, blacks have many conservative values that align better with the Republican Party, and the Democratic Party is not as good for the black community as the black community thinks it is (Cozier, 2011; Wiggins, 2011).

Rather than present substantive arguments for why blacks should consider voting Republican, these billboards made the case for voting Republican by appealing to a sense of social propriety in the black community: Martin Luther King, Jr. voted Republican – *so should you*. Hip young black people are voting Republican – *so should you*. The unstated premise of this rhetorical approach is that many black voters currently do not vote Republican because it is not socially acceptable.

The question of why blacks vote Democratic is, in part, a question of social influence. In political science, the study of social influence involves understanding the relationship between individuals and groups and how individuals influence each other within these groups. To find out what role social influence plays in mass politics, researchers have a variety of methods and models at their disposal. These methods and models have led researchers to conclude that social influence is important in politics. For example, through casual conversations, individuals may learn about others' attitudes toward specific issues and may come to adopt those feelings themselves (Mutz, 2002; McClurg, 2006). But the idea that social influence is important in politics is hardly confined to scholarly research. For example, politicians often identify themselves as members of various social groups, implicitly assuming that members of a particular group are affected by what other members of that group think. Before the advent of modern methods and models, Aristotle believed that the fundamental unit in politics was not the individual, but the "partnership," or relationship between individuals, such as husband-wife, parent-child, and ruler-subject (Aristotle, 1995 [350 BC]).

An important implication of the idea that social influence is important in politics is that individuals tend to be *interdependent* in their political behavior and attitudes. Ignoring this interdependence has important methodological consequences for political science research. First, ignoring interdependence in political behavior and attitudes may lead to a misspecification of both individual-level and group-level factors that underlie political behavior. According to prominent social influence researcher Robert Huckfeldt (2007a),

> If various forms of political behavior are, in turn, contingent on an individual's location within networks of political communication, the likelihood of engaging in a behavior – holding an opinion, voting for a candidate, putting up a political yard sign – is apt to vary across these various contextual units of aggregation. Thus, aggregate analyses that ignore important patterns of interdependence enhance the risk of producing ecological fallacies.[1]

Second, ignoring interdependence in political behavior and attitudes may lead to a misspecification of the relationship between individual-level factors and individual-level outcomes. In contrast to the ecological fallacy, this atomistic misspecification may lead to inaccurate conclusions based on individual-level data; Huckfeldt called this an "individualistic

[1] An ecological fallacy refers to the inaccurate conclusions that arise from using aggregate (*i.e.*, "ecological") data to infer individual-level relationships when individual-level data are not available (see King, 1997).

fallacy." According to Huckfeldt (2007a), both methodological consequences are based on "the stated or unstated assumption that individual characteristics and attributes translate directly into likelihoods of opinions and behaviors independently of the networks and contexts within which individuals are embedded."

Ecological and individualistic fallacies point to the importance of basing methodological assumptions on realistic and substantively meaningful observations of human sociability. Of course, individualistic or aggregate models of political behavior and attitudes might be defended as mere abstractions that are, by nature, coarse approximations of real, complex human actions. But even mere abstractions should be no coarser than they need to be.

RESEARCH QUESTION

That politics is inherently social is hardly controversial. But the issue of how to study social influence, in light of the fact that politics is inherently social, is more so. Accordingly, the research question at the heart of this dissertation is, "How should we study social influence?"

Political scientists have conventionally examined social influence through the analysis of variance (ANOVA) framework, social network analysis, contextual analysis,[2] and linear regression analysis. While these tools are useful and have provided important insights into the nature of social influence in politics and other areas of study, they have important methodological issues (as we will see) and often fall short of providing meaningful information on the social influence process under study. To enhance the study of social

[2] A special case of hierarchical linear modeling

influence, this dissertation advocates and uses spatial econometrics as an alternative, more meaningful approach.

SOCIAL INFLUENCE AND THE PROBLEM OF RECIPROCAL CAUSATION

One of the most pressing issues in the study of social influence is the problem of reciprocal causation. In political science, the study of social influence involves understanding how, why, and to what extent individuals are influenced by their social environments. The problem of reciprocal causation is an identification problem regarding the direction and size of influence between an individual and his social environment. McClurg (2010a) provides the following example of reciprocal causation in a study of boss-employee relationships and their impact on politics:

> If the employee donates money to a specific presidential candidate, as does her boss, is it because the boss has persuaded her to make that donation in lieu of facing professional setbacks? Such an argument is perfectly sensible and gets raised frequently in discussions of campaign reform. However, it is equally possible that it is the employee, anxious to demonstrate her political perspicacity, explains to her boss the importance of making the contribution in order to advance the company's interests. It is also possible that *both* processes are acting simultaneously to produce the behavior of interest.

In general, the problem of reciprocal causation is the problem of isolating the effect of a social environment on an individual when it is possible that the individual might also have an effect on his social environment; that is, the effect flows not only from the environment to the individual but also from the individual to the environment. Unfortunately, conventional methods for studying social influence, such as contextual analysis, often fails to provide a satisfactory answer to the problem. In his critique of contextual analysis, Blalock (1984) wrote, "[T]o the degree that the causal processes [differ], it would also be difficult, if not impossible, to assess the direction of causation between the individual- and the contextual-level variables."

The problem of reciprocal causation is an important one because it calls into question the internal validity of social influence studies.[3] Despite the volume of research on social influence, researchers have not provided definitive evidence of a causal relationship between interactions in social environments and political behaviors and attitudes. Since the direction of causation is difficult to determine, explanations of social influence processes are often incomplete and unsatisfactory (see Klofstad, 2007 for further discussion).

Researchers have attempted to address the problem of reciprocal causation in several ways. One approach is to employ systems of equations involving variables that independently predict values of the dependent variable and characteristics of the social environment under study. According to McClurg (2010a), however, these statistical solutions are "frequently intractable and impractical" and "often provide only imprecise and rough estimates about casual effects." Another approach is to use creative research designs to address internal validity explicitly, such as field experiments and natural experiments (Klofstad, 2007; Nickerson, 2008). Unfortunately, such experiments have well-known problems of their own, such as the problems of external validity, generalizability, and feasibility. The shortcomings of these and other methods led McClurg (2010a) to express pessimism regarding a solution to the problem of reciprocal causation:

> [W]hile each approach is helpful, none of them holds out the promise for solving these problems "once and for good" in the field of social communications. This again suggests to me that healthy progress in the subfield of social communications depends deeply on the use of multiple methods and the slow accumulation of evidence, rather [than] a single set of methodological solutions to causal validity issues.

[3] According to Huckfeldt and Sprague (1991), the problem of reciprocal causation may be important in theory but not necessarily in practice. Distinguishing between egos as the centers of social networks and alters as members of those social networks, they found that very few alters choose egos as discussion partners among nonrelatives, which provided little evidence of actual reciprocity. However, their result is confined to non-relatives and the larger theoretical issue of validity remains.

While this dissertation does not solve the problem of reciprocal causation "once and for good," it does present a methodological approach that directly addresses the issue of reciprocal causation and has the potential to aid "healthy progress" in research on social influence in politics.

SOCIAL INFLUENCE AND SOCIAL DISTANCE

A fundamental prerequisite for social influence is social interaction, since individuals must have the opportunity to interact with one another if social influence is to take place. The likelihood of social interaction depends, in large part, on *social distance*.

Social distance is a general measure of similarity between individuals in a particular social environment. Social distance can be based on geography, demographics, ideology, or other characteristics of a social environment where social influence might take place. The idea is that the more similar two individuals are, the less social distance there is between them, and the more likely they are to influence each other. For example, there is less social distance between individuals who live in the same city, compared with individuals who live in separate cities. There is less social distance between individuals who are in the same income bracket, compared with individuals who fall into different income brackets. There is less social distance between individuals who belong to the same political party and greater social distance between individuals who belong to different political parties. Accordingly, individuals who live in the same city, who fall in the same income bracket, and/or belong to the same political party are more likely to interact and mutually influence each other than individuals who are more socially distant.

Conventional approaches to the study of social influence tend to take a binary view of social distance, which leads to a binary view of social influence. For example, a researcher might model an individual's vote choice as a function of the mean political ideology of each individual's city. This means that the ideologies of fellow city dwellers might have an impact on each individual's vote choice, but the ideologies of those who live in other cities, even in cities that are adjacent or very close to the individual's city, are assumed to have no impact at all. This assumption regarding the nature of social influence may be overly restrictive, especially in cases where cities are relatively small and individuals are very mobile. In such cases, there may be potential for a lot of interaction and hence social influence among individuals from *neighboring* cities. Furthermore, the amount of interaction and social influence might be closely related to the distance between the cities; there might be a lot of interaction and influence among individuals in the same city, relatively less interaction and influence among individuals in neighboring cities, and no interaction or influence among individuals who live in cities that are far apart.[4] Conventional approaches, such as contextual analysis, do not allow researchers to scale social influence according to distance in this manner. By treating social influence as an all-or-nothing phenomenon in all circumstances, researchers must necessarily assume that there can be social influence only between individuals who live in the same city. More generally, researchers must assume that social influence is limited to individuals within discrete categories of interest.

One might argue that conventional methods do in fact account for social influence (to a limited extent) though the use of control variables in linear regression models. In political science research, researchers may wish to account for social environments, even if they are not interested in modeling social influence directly. In such a circumstance, a researcher might include several demographic variables as control variables in a linear regression model. For example, race, sex, social class, and educational attainment might be

[4] The non-binary nature of social influence was suggested in the work of Berelson et al. (1954), who noted that the stronger the relationship of individual Catholics with their local church, the stronger the effect of the group norm on individual political preferences, compared with the community norm in Elmira.

included as independent variables in a model of individual attitudes toward government spending. These independent variables can be interpreted as indicators of group interests, such as women's interests, Hispanic interests, or business class interests. This interpretation is based on the social-psychological thinking that is common in political science research. As Achen (1992) noted, "Group identification was said to structure individuals' thinking, and the stronger the identification, the stronger the effect." Accordingly, group identifiers, such as party identification and demographics can serve as controls for group loyalties that might impact political behaviors and attitudes.

Unfortunately, using control variables to account for social influence has important limitations. First, indiscriminately using control variables to account for important but auxiliary factors may produce misleading coefficient estimates in the presence of unrecognized nonlinearity (see Achen, 2002; Achen, 2005). Thus, while the practice of using many control variables may be widespread and convenient, it may be methodologically suspect. Second, the assumptions may be unrealistic with regards to social influence. The impact of control variables is assumed to be linear and additive, and each causal effect presumably operates independently of the values of other variables. This means that the researcher would have to treat certain control variables, such as race and class, separately, even if they are related in a social environment. While researchers might overcome this problem by interacting all the control variables, such a solution might lead to an unacceptable loss of degrees of freedom. Even if the loss of degrees of freedom is acceptable, a linear regression model with control variables and interacting control variables cannot adequately account for social distance because each respondent would be identified with a particular combination of control variables and no others and because there is no guarantee that greater levels of similarity would correspond with greater levels of social influence. This means, for example, that researchers would be able to assess the effects of being working class and Hispanic on an individual's vote choice, but not the effects of other classes or other races, even if they are relevant.

THE CENTRAL THEORETICAL CONTRIBUTION

This dissertation takes the view that reciprocal causation and social distance can be addressed methodologically through the use of spatial econometrics.

Previous studies have failed to address the problem of reciprocal causation because researchers have often used models and methods that inherently assume that social influence flows only in one direction, leaving the problem of reciprocal causation entirely to the discretion of the researcher. As Huckfeldt and Sprague (1991) point out, however, "[I]nfluence flows in two directions . . . ." Rather than use models that assume a uni-directional process of influence, researchers should instead use models that actually incorporate mutual influence and social distance among individuals. In this dissertation, we will see that the field of spatial econometrics offers an attractive set of models that do exactly this.

Spatial econometrics is a set of tools for modeling mutual influence among interacting individuals (or other units). It allows researchers to find out, all things being equal, whether an environment (*i.e.*, a "neighborhood" in spatial econometric jargon) exerts any influence on an individual, after accounting for the individual's role in that environment. Doing this involves using a spatial regression model (*i.e.*, a spatial lag model) that represents the social influence process as a dynamic diffusion process in which individuals mutually influence each other in specific types of neighborhoods. Furthermore, such a model features the size and direction of social influence as a parameter, which means that it can be estimated, rather than relegated to theoretical exposition and speculation. As this dissertation will show, this type of model is much more representative of social influence as *mutual* influence and explicitly accounts for the interdependencies among individuals in their social environments.

Indeed, spatial econometrics is especially apt for estimating the effects that others' behavior or attitudes have on the behavior or attitudes of individuals. While conventional methods, such as contextual analysis, can only describe a group's effect on a member, spatial econometrics can show what effect the member has on the group. Take, for example, the decision of baseball player Albert Pujols to sign with the Los Angeles Angels of Anaheim. Contextual analysis can describe how joining the Angels changes Mr. Pujols, but it cannot describe how changes in Mr. Pujols might affect the Angels. Suppose we are interested in the average amount of time baseball players spend practicing base running, and suppose that there is a culture of base running among the Angels so that each team member practices more base running, compared with being a member of a different team. If Mr. Pujols loses weight because of this running, causing him to practice even more base running, what effect would his extra practice have on the amount of time that his team members spend base running? Would Mr. Pujols' weight loss indirectly affect the average amount of base running practice among the Angels? As this dissertation shows, spatial econometrics can answer these kinds of questions.

Furthermore, spatial econometrics allows researchers to account for social distance when modeling social influence. Spatial regression models feature what is called a spatial weights matrix, which is a square matrix that represents the relationships between the respondents in a sample. These relationships can be as simple as common membership in a group (such as a church), which can be represented in a binary manner; or the relationships can be more nuanced, such as language similarity (which can be represented on a continuous scale). As we will see, the spatial weights matrix is a powerful and flexible way of representing social environments of interest.

BRIDGING SOCIAL INFLUENCE AND SPATIAL ECONOMETRICS

While spatial econometrics originated in the study of geography, this dissertation seeks to broaden its use to the study of social influence in politics. This effort is in line with the increasing interest in and use of spatial econometric tools in political science over the past decade. Researchers have used spatial econometric models to study such diverse phenomena as the diffusion of policies and innovations (Werck et al., 2006; Woods, 2006; Couglin et al., 2007; Volden et al., 2008), the role of interdependencies in international relations (Gleditsch and Ward, 2000; Gleditsch, 2002; Ward and Gleditsch, 2002; Gleditsch, 2007) and political economy (Simmons and Elkins, 2004; Simmons et al., 2006; Beck et al., 2006), neighborhood effects on national identity (Lin et al., 2006), and neighborhood effects on campaign contributions (Cho, 2003).

Expanding the use of spatial econometrics is not a straightforward task. Spatial econometrics is a set of tools that originated in the study of geography; its concepts, models, and methods require explanation as to how they can be applied to the study of social influence in politics. This dissertation provides this explanation in several ways.

First, it explains how social influence can be understood as a form of spatial dependence and can be analyzed as spatial autocorrelation. A primary concern in spatial econometrics, spatial dependence is the relationship between observations and the locations of those observations.[5] Anselin (1988) defines spatial dependence as follows: "In general

[5] Spatial dependence can be compared with serial correlation in time-series analysis. In the former case, observations are lagged over a given subset of other observations, whereas in the latter, observations are lagged over a given time period. While the two concepts can be considered analogous, Anselin (1998) cautions:

terms, spatial dependence can be considered to be the existence of a functional relationship between what happens at one point in space and what happens elsewhere." He points out that where human beings are located and how far apart they are from each other are important elements in structuring human behavior. This dissertation is an attempt to bring this insight to the study of social influence and address the problem of reciprocal causation.

Second, this dissertation shows how spatial econometrics can allow researchers to work with more flexible definitions of social context. It can deal explicitly with interdependencies among groups that are not necessarily geographically-based. A primary contribution of this dissertation lies in re-conceptualizing social context as space and social influence as functions of distance in this space. This space need not be geographically-defined; it need not be limited to census tracts or state boundaries. Taking Beck et al.'s (2006) suggestion that "space is more than geography," this dissertation explores this idea by translating different types of social contexts and social influence into spatial terms. Specifically, this dissertation conceptualizes three different types of spaces – *i.e.*, geographical, demographic, and ideological – and looks at whether and to what extent individuals mutually influence each other's political attitudes in these spaces.

Third, this dissertation provides in-depth discussions of spatial econometric models and conventional models of social influence. These discussions address how and why spatial econometrics can provide information on the phenomenon of social influence that other methods cannot. The purpose of these discussions is to demonstrate the theoretical

---

Upon closer examination, it often follows that standard econometric results from time series analysis do not carry over in a straightforward way to spatial dependence in cross-sectional samples. This is primarily a result of the multi-directional nature of dependence in space, which, as opposed to a clear one-directional situation in time, precludes the application of many simplifying results and necessitates the use of a different methodological framework.

advantages of spatial econometrics and why social influence researchers should include it in their methodological toolbox.

Finally, this dissertation provides three illustrative examples of how econometric models can be used to understand the role of social influence in politics. In three empirical chapters, this dissertation assesses whether and to what extent individuals who are geographically, demographically, and ideologically similar mutually influence each other's attitudes toward political issues such as government spending and access to abortion. The models used in these assessments incorporate social contexts that are not limited to geography; represent social influence as a dynamic process of mutual influence among interdependent individuals; and allow for the estimation of a social environment's influence on individuals, after accounting for reciprocal causation and individual-level factors.

ORGANIZATION OF THE DISSERTATION

In this introduction, I have discussed the major themes, research question, and scholarly contribution of this dissertation.

In chapter 2, I discuss the key concepts, theories, and propositions pertaining to the role of social influence in mass politics and the spatial econometric framework. I argue that a social context can be reconceptualized as a space for social interaction, that individuals who might be subjected to social influence occupy particular points in such a space, that the extent of social influence is related to the "distances" between such individuals situated in the space, and that similarly-situated individuals comprise a kind of spatial "neighborhood," or group, in which each individual influences and is influenced by his spatial neighbors' political behaviors and attitudes. Then, I explain why social influence must be understood as mutual influence and argue that models of social influence must account for the endogeneity in the dependent variable of interest. Next, I introduce spatial econometrics and show how the concepts, theories, and propositions of social influence come together in spatial regression models of social influence. Finally, I introduce existing models of social influence, compare them to spatial regression models, and show how the latter models better represent the social influence mechanism.

Chapter 3 provides an in-depth comparison of spatial regression and linear regression models to address directly the issues of social influence and social distance. In this chapter, I examine ANOVA, simple linear regression, and multiple linear regression models; discuss their relationship to spatial regression models; and show how spatial regression models are more substantively meaningful and statistically informative for the study of social influence than linear regression models.

Chapter 4 describes the research design of the dissertation. In this chapter, I identify and discuss the unit of analysis, the data source, measurement issues, estimation methods, and testable hypotheses employed in this dissertation's empirical analyses of the role of social influence on individual attitudes toward political issues. Furthermore, I also discuss Galton's Problem, which has to do with the enormous difficulty of separating the effects of mutual influence from common shocks, and how this dissertation deals with this issue.

Chapters 5, 6, and 7 are empirical applications of spatial econometrics to the question of whether and to what extent social influence impacts the political attitudes of individuals. In these three chapters, I provide more detailed accounts of how social influence on political attitudes might take place in geographical space, Blau space, and latent ideological space. I put into practice the idea that the dimensions of these spaces can be understood not as control variables, as they have often have been used, but as structures underlying the clustering of political attitudes. Implementing the research design discussed in chapter 4, I construct spatial weights matrices based on geographical, Blau, and ideological distances and use these matrices in spatial regression models of social influence to find out whether and to what extent individuals who are similarly situated in geographical, Blau, and ideological space mutually influence each other's political attitudes.

Chapter 8 concludes the dissertation. It highlights the main issues and findings of this dissertation, and discusses their implications for the study of social influence in politics. Finally, it identifies avenues for further research.

CHAPTER 2: THEORIES AND MODELS OF SOCIAL INFLUENCE

As individuals make up their minds about political issues or decide whether to act politically, it is plausible that they take into account the views of family members, friends, colleagues, and neighbors. These family members, friends, colleagues, and neighbors are part of each individual's social context. But depending on the researcher's subject of study, an individual's relevant social context may be narrower or broader than these. An individual might account only for the views of his family when deciding whether to vote for a property tax hike; or he might account for the (perceived) views of several countries in deciding whether to support or oppose his country's decision to go to war. Thus, in the study of social influence, it is important to specify the relevant social context that circumscribes the subject of study. This important specification bridges substantive and methodological concerns; it encourages more precise theories of social influence, constrains the set of possible social influence mechanisms under consideration, and helps determine which models and what data to use for the empirical study of social influence.

Accordingly, this chapter begins with a discussion of different definitions of social context and their consequences for understanding social influence. I argue that the increasing number of competing definitions has led to greater confusion regarding the mechanisms of social influence. To remedy this problem, I suggest a re-conceptualization of a social context as an abstract, multidimensional space. This idea of social context as space is general yet useful; it can subsume existing notions of social context while providing a clearer framework for positing mechanisms for social influence. Next, I give three examples of social contexts and discuss how thinking of them as spaces helps to show how they might shape political attitudes.

Along the way, I also clarify the major concepts supporting the idea of social context as space, such as social space, social distance, and neighborhood. Using these concepts, I argue that social influence should be understood as *mutual* influence among individuals who occupy particular points in a space, that the extent of social influence is related to the "distances" between such individuals situated in the space, and that similarly-situated individuals comprise a kind of spatial "neighborhood," or group, in which each individual influences and is influenced by his spatial neighbors' political behaviors and attitudes. Next, I show how this framework can be used with spatial econometric tools for empirical analyses of whether political attitudes are subject to social influence. To do this, I provide a brief introduction to spatial econometrics and show how the concepts, theories, and propositions of social influence come together in spatial regression models of social influence. Finally, I introduce existing models of social influence, compare them to spatial regression models, and show how the latter models better represent the social influence mechanism.

THE EMERGENCE OF "SOCIAL CONTEXT"

"Social context"[6] refers to the characteristics of social aggregations such as neighborhoods, discussion groups, peers, small groups, and social networks. The relevance of social context to political behavior and attitudes has long been established in social science. For example, political discussions in social networks affect how individuals view and

[6] Also called "social environment" or "social structure."

participate in elections (Huckfeldt and Sprague 1991, 1995; Kenny 1992, 1994). Lake and Huckfeldt (1998) found that more frequent political discussion in individuals' social networks correlates with higher levels of political participation. Social aggregations also affect individuals' attitudes toward specific issues (Binder et al., 2009), cause individuals to hold more extreme attitudes (Sunstein, 2000), decrease ambivalence about political decisions (McClurg, 2007), change initial opinions about politics (Luskin et al., 2002), and increase political participation (McClurg, 2007).

The study of social contexts and their consequences emerged from the work of prominent sociologists such as Durkheim, Berelson, Lazarsfeld, and Blau. Observing the rise of mass surveys in social science research in the 1950s, Columbia sociologists such as Berelson and Lazarsfeld worried about the possible marginalization of social interactions and complexities (that in-depth studies of group dynamics could better explain), and about the problems with inferring individual behavior from aggregate data (the "ecological fallacy"). Their solution to both problems was the analysis of "contextual effects," which were also known as "structural effects" or "compositional effects." The analysis of contextual effects involves obtaining measures on individuals, aggregating the measures based on group membership to obtain group attributes, and then determining the effect of a group attribute on some individual-level outcome variable when the corresponding individual characteristic is held constant. The effects of the group-level independent variables are the "contextual effects" that presumably arose from social interaction within a social context. Blau (1957) explains:

> The general principle is that if ego's X affects not only ego's Y but also alter's Y, a [contextual effect] will be observed, which means that the distribution of X in a group is related to Y even though the individual's X is held constant. Such a finding indicates that the network of relations in the group with respect to X influences Y.

This general approach uncovered a number of interesting social phenomena. For example, Berelson, Lazarsfeld, and McPhee (1954) found that the higher the proportion of Republicans in a group of friends, the higher the proportion of non-Republicans in that group who vote Republican. Lipset, Trow, and Coleman (1956) found that political interest is higher in social contexts with higher levels of political consensus, regardless of political preference. Davis, Spaeth, and Huson (1961) later formalized the general approach outlined by Blau and others by developing techniques for identifying contextual effects and a typology of different kinds of contextual effects.

In the 1980s, Robert Huckfeldt and his colleagues took the sociologists' analysis of contextual effects and adapted it to the study of political psychology and public opinion by collecting egocentric network data and using contextual analysis. Egocentric data collection[7] was a way of incorporating social context data into random sample surveys. Contextual analysis provided the statistical methods and models for parsing individual-level and group-level effects within a regression framework. In order to apply both of these tools, however, Huckfeldt and his colleagues found that they needed a major redefinition of social context, which had been too conceptually narrow for their needs.

Prior to these research efforts, "social context" was not a major conceptual issue because it was understood simply to refer to the observable characteristics of clearly-defined groups. Early sociological studies of contextual effects involved clearly-defined groups such as assembly line workers, counties, the military, religious sects, and countries, which were chosen prior to the actual research effort. This approach carried over to early studies of contextual effects in political science as well, when researchers focused on clearly-defined

[7] In egocentric data collection, researchers collect data on individuals and their associates in a multi-step process. First, they conduct the initial interviews with randomly-selected individuals from a population of interest. Next, the researchers ask those individuals to name other individuals with whom they have discussions. Finally, the researchers carry out a second set of interviews with the named discussants.

groups based on election districts (Tingsten, 1963), towns (Segal and Meyer, 1974), mining and resort areas (Butler and Stokes, 1974), and counties (Miller, 1956; Putnam 1966).[8]

But over the course of their research, Huckfeldt and his colleagues found that they needed a broader, more generally applicable definition of social context in order to take advantage of their new data and methods. Accordingly, Huckfeldt took the idea of a "neighborhood," which conventionally refers to a cohesive geographical community, and reworked it into a more abstract concept that included, but was not limited to, the conventional understanding. This broader definition can be found throughout much of Huckfeldt's research (see Huckfeldt, 1979; Huckfeldt, 1983a; Huckfeldt, 1983b; Huckfeldt, 1986). A typical definition of "neighborhood" is as follows:

> The neighborhood, as it is conceived here, refers to the shared geographical locale of a residential grouping, and the neighborhood social context refers to the population composition of the people who live in the neighborhood.
>
> This neighborhood definition subsumes the cohesive community but does not depend upon it. The definition is designed to capture the inescapable social relations of any geographically based social collectivity. Thus, the neighborhood is of interest as a structural factor in the lives of its residents, rather than as a well-articulated social organization. (Huckfeldt, 1986)

The "well-articulated social organization" refers to clearly-defined groups such as churches, assembly plants, and other subjects of early sociological studies of contextual effects. By expanding the idea of a "neighborhood" from these kinds of groups to "any geographically based social collectivity," Huckfeldt allowed the idea of social context to be more general. For example, under the narrower, sociological definition, an assembly line worker's social context is restricted to the relationships with his fellow workers and supervisors; under

[8] Some of these early studies indicated a need for a broader view of social context. Putnam (1966) and others used counties as the contextual units but not in the same spirit as the early sociological studies. Instead, they assumed that the attributes of counties are good indicators of the attributes of an individual's immediate social context.

Huckfeldt's broader definition of "neighborhood," that same assembly line worker's social context can include the residents of his hometown – even the ones he has never actually met or interacted with. This broader conception of social context provided theoretical justification for using broad indicators of a person's social environment, such as the average income level in a county or percentage of blue collar workers in a town.

Huckfeldt's generalization of the concept of neighborhood and, by extension, of social context, allowed for the wider application of contextual analysis to the study of mass politics. This generalization brought increasing attention to the question of how social influence actually works.

THE ISSUE OF SOCIAL INTERACTION

Interactions among individuals in well-defined organizations (like assembly plants) are apparent, while interactions among individuals in broadly-defined social environments (like Huckfeldt's neighborhoods) are much less explicit. Accordingly, the broadening of "social context" must be accompanied by a different understanding of what social interactions in particular social contexts look like. Huckfeldt's definition of neighborhood assumes that mere geographical proximity leads to "inescapable social relations." These relations do not depend on some cohesive sense of community or friendship among members, or on whether those members choose to engage in those relations (hence "inescapable"), because the neighborhood imposes associational constraints and opportunities on each individual: "[T]he social content of social networks is not solely a function of either the social context or individual choice; *it is the complex product of individual preference operating within the boundaries of a social context*." (Huckfeldt, 1986) Whereas Blau argued that a "structural effect" is the effect due to social interaction within a well-defined group, Huckfeldt modified this idea so that a "contextual effect" is the effect to due to the *constraints* on social interaction within a loosely-defined group.

But Huckfeldt's idea of social contexts as social constraints raised serious questions about the nature of the social interactions themselves. How does the presence of political lawn signs and bumper stickers in a neighborhood actually affect how individuals think about politics? How do casual or accidental encounters with near-strangers actually facilitate social influence?

Perhaps in recognition of this problem, researchers began distinguishing between interpersonal and impersonal contexts. Huckfeldt and his colleagues' theory that social contexts constrain social interactions did not explain how those social interactions actually worked. Based on their theory of social context as associational constraint, social influence is only an *indirect* effect of one's environment. But what about the *direct* forms of social influence that actually take place within those contextually-constrained social relationships?

To understand social influence within associations, some researchers turned to social network studies to understand how social influence works, given that actual interaction can be assumed. McClurg (2010a) provides the following definition of a "social network":

> The key definitional element of a social network is the presence of identifiable relationships between people where conversations create opportunities for the transfer of politically relevant information, such as pertinent political facts, general perspective on politics, political norms and mores, and so on.

Social networks are *interpersonal contexts* that involve intimate social interactions between individuals, such as those among family and friends. Interpersonal contexts are the concern of social network studies, which examine the relationship between individuals' political behavior and the attributes of their personal network of friends, families, and colleagues.

Impersonal contexts involve casual, day-to-day encounters between individuals who are less intimately related to each other. Impersonal contexts are the concern of aggregate studies, which examine the relationship between individuals' political behavior and the attributes of the general social context (*e.g.,* election districts, ZIP code area) in which those individuals are located. The importance of impersonal contexts gained recognition with the work of Mutz (1992), who brought attention to the mass media's role in facilitating "impersonal influence."

> The power of these collective representations flows from the sheer numbers they claim to represent rather than from their specific identities. It is in this sense that impersonal influence may be viewed as a fundamentally different social influence process with fundamentally different mechanisms of influence from conformity or group identification.

The split between aggregate studies and social network studies necessarily raised the question of what social interaction and social influence really mean in these two areas. In aggregate studies, explanations of social interaction and social influence tend to follow Huckfeldt's view that social contexts constrain associational opportunities. But in social network studies, explanations of social interaction and social influence tend to follow Blau's view that social contexts consist of actual social interactions that shape behavior. How can we reconcile these two views to understand the overall relationship between social context and social interaction? In the following section, I suggest a possible reconciliation.

SOCIAL SPACES AND NEIGHBORHOODS: AN ALTERNATIVE FRAMEWORK

> [A] contextual explanation for political behavior emphasizes the interdependence of political choices made by individuals who share a common social space.
>
> - Robert Huckfeldt, 1983

A first step toward reconciling Blau and Huckfeldt is to climb up the ladder of generality. Both Blau and Huckfeldt would agree that social context is a kind of structure containing diverse and numerous human behaviors. While they would disagree on the specific form of the structure itself, they would agree that such a structure could be characterized by certain characteristics, such as race, political ideology, or income level. To maintain Blau and Huckfeldt's commonalities, we can think of social context as an abstract "social space" in which individuals inhabit and potentially or actually interact. Individuals "inhabit" points in the social space in the sense that they have certain roles that belong to the structure. In the case of Blau, for example, the assembly plant is not just a physical place; it also represents an abstract space where individuals have roles as assembly workers, foremen, or supervisors. One could describe an assembly plant as a two-dimensional social space where one dimension is the formal ability to exercise control over others and the other dimension is the informal propensity to demand work or assistance from others. An assembly worker's position would be low on the first dimension but middling on the second, while a foreman would be middling on the first, but high on the second and a supervisor would be high on both dimensions. In the case of Huckfeldt, the census tract is not just an arbitrarily-defined geographical area for census-taking; it can also be viewed as a two-dimensional space where individuals move around in their day-to-day lives and potentially encounter each other.

By thinking of social contexts as social spaces, one can go beyond the geographically-defined social collectivities that Huckfeldt and his colleagues studied. Geography may have played a huge role in constraining social relations back in the 1970s and 1980s, but there are reasons to suspect that this role has diminished. Technological advances have allowed for instantaneous communication, which allow social relations to traverse geographical boundaries and to flourish despite the absence of physical, day-to-day, face-to-face interaction. In fact, the Millennial generation, which has grown up with these technological advances, regard tweeting, texting, and blogging as everyday parts of their social lives (Taylor and Keeter, 2010). Political scientists have noted the lessening of geographical constraints. "Geographical contexts, such as neighborhoods, are increasingly less important as spaces for social interaction," observed McClurg (2006). Second, individuals are less tied down to their jobs and families than they used to be (Taylor and Keeter, 2010). To the extent that jobs and families determine geographical location, this trend suggests that individuals may not be as strongly attached to the geographical location of where they live as strongly as they used to be.

Defining the social context is important because it determines what the group-level variables are. If the social context of interest is the neighborhood, then group-level variables might include the percentage of Republicans or the average level of interest in political campaigns. But if geography is only part of an individual's social context, then acquiring group-level variables in this way may be too restrictive. In contrast, a social space can be broadly defined by any number of characteristics, such as occupation, age group, and culture, and can comprise any number of dimensions.

Since all individuals can be viewed as inhabitants of a social space, individuals who share the same or similar positions in the space can be considered "spatial neighbors." For example, all assemblers in the assembly plant can be considered neighbors, since they have similar positions on the two dimensions of social space defined by the assembly plant. Inhabitants of the same census tract can be considered neighbors since they are geographically proximate. These spatial neighbors comprise a "spatial neighborhood," and the set of all spatial neighborhoods comprise the social space. This definition of neighborhood is much broader than the usual meaning of the term, which refers to an actual place where people live. It is also broader than the sense used by geographers, for whom a "neighborhood" is simply a geographically-proximate area without cultural meaning. But this definition is flexible enough to include both usages, as different as they are.

In a social space, individuals potentially or actually interact according to some rule.[9] This rule determines to what extent individuals interact with each other. In the assembly plant example, the rule might be that the more similar jobs are in terms of status, the more interaction there would be. There would be a lot of interaction among the assemblers but less interaction between assemblers and foremen, and even less interaction between assemblers and supervisors. Another rule might be that reporting requirements determine the frequency of interaction. Under this rule, there would be a lot of interaction between assemblers and foremen and between foremen and supervisors but not a lot of interaction among assemblers or among foremen or among supervisors. In the census tract example, the rule might be that individuals who live in the same census tract have a greater chance of meeting each other than individuals who do not live in the same census tract. Another rule might be that individuals who live in the same census tract as well as adjacent census tracts have the same chance of meeting each other but individuals who do not live in the same or adjacent census tract have no chance of meeting. Rules of interaction can be as specific as the assembly plant example or as broad as the census tract example.

For the most part, identifying these rules depends on the amount of specificity desired or attainable by the researcher. However, these rules must be based on substantive mechanisms that represent actual processes of interaction and influence and link back to the observed contextual effect. According to Erbring and Young, "[C]ontextual effects must be conceptualized as a consequence of processes of interaction among individuals in a social

[9] Similarly, Douglass North (1991) defines institutions as "humanly devised constraints that structure political, economic and social interaction. They consist of both informal constraints (sanctions, taboos, customs, traditions, and codes of conduct), and formal rules (constitutions, laws, property rights)." However, my use of the term "rule" in this dissertation is much more general and is not limited to institutions.

network." They suggested two broad classes of interactions: Actual face-to-face contact between pairs of individuals or symbolic categorical relations shared among all individuals in a given group. *Face-to-face interaction* involves processes of contagion, diffusion, persuasion, assimilation, conformity, consensus, and contrast, while *symbolic interaction* involves processes of comparison, competition, emulation, identification, facilitation, and inhibition. According to Erbring and Young, researchers must reference either type of social interaction, lest "the notion of contextual effects tends to become theoretically vacuous."

THREE TYPES OF SOCIAL SPACES AND DISTANCES

In the study of political attitudes, demographics, ideology, and geography are important sources of social influence on how individuals think about political issues. First, demographics are important indicators of group loyalties, which help structure individuals' thinking about politics. For this reason, demographic variables such as race and sex are often included in linear regression models of political behaviors and attitudes (Achen, 1992).

Second, ideology serves as constraint on how individuals think about politics, not only in the individual sense, but also in a social sense. While researchers tend to focus on ideology as an *individual-level* characteristic that may or may not impact individual political behavior or attitudes (Campbell et al., 1960; Converse, 1964; Kinder, 1998; Ansolabehere, 2008), there is research suggesting that there are *social* aspects to ideology as well. According to Elkins and Simmons (2005),

> [I]mitating similar individuals is one of the simplest and most effective cognitive heuristics in the calculations of utilities. Actors negotiating a complex set of political choices regard the actions of actors with perceived common interests as a useful guide to their own behavior.

These "useful guides" can be identified through visible characteristics such as geographical and cultural proximity. In a study of policy diffusion, Sugiyama (2008) identified ideological proximity as another useful guide, since policy makers are more likely to emulate the policies of ideologically-similar municipalities, even in the absence of partisan directives.

Third, geography has been shown to be an important factor in how individuals think and act in politics because the communities in which individuals live and work are often circumscribed by geographical boundaries and individuals are susceptible to community standards and norms (Campbell, 1960; Rodden, 2010).

In this section, I reframe demographics, ideology, and geographical location as Blau space, ideological space, and geographical space, and identify rules for interaction in each space. These spaces are based on variables that social scientists often use as control or explanatory variables. But here, these control variables are reconceptualized as dimensions of social spaces for social influence.

Blau space and Blau distance

Coined by sociologist Miller McPherson, Blau space is "the *k*-dimensional system generated by regarding the sociodemographic variables as dimensions, rather than as variables." (McPherson, 2004) In Blau space, each individual has a position described by *k* coordinates, and distances in the coordinate system define the relationships among the points. Given a three-dimensional Blau space defined by income, race, and sex, for example, a middle-class black woman and a poor white man are two distinct points in the space. The distance between these two points is relatively large; in contrast, the distance between a middle-class black woman and a middle-class black man is relatively small. Thus, demographic characteristics can be understood not as variables but as a structure underlying the clustering of political attitudes. This structural understanding of social characteristics is McPherson's homage to Blau, for whom the space is named.

Homophily is the organizing principle in Blau space. The intuition underlying the principle of homophily is that like attracts like; it is the sociologists' refinement of the popular notion that "birds of a feather flock together." Formally, according to McPherson et al. (2001), "Homophily is the principle that a contact between similar people occurs at a higher rate than among dissimilar people." This implies that friends, family, and other personal networks tend to be relatively homogenous when it comes to many characteristics.

Social scientists have found that the characteristics that matter the most are demographics: race/ethnicity, age, religion, education, occupation, and gender (McPherson, et al., 2001). Homogeneity in terms of these characteristics consistently shows up in survey respondents' descriptions of those with whom they discuss politics or other important matters. Using network data from the 1985 General Social Survey, Marsden (1988) found that respondents' discussion networks tended to share similar demographic attributes. This, he argues, is evidence that similarity generally breeds stronger ties among individuals than does dissimilarity. In a subsequent study, McPherson et al. (2006) compared network data from the 1985 General Social Survey with network data from the 2004 General Social Survey and found that the heterogeneity of confidantes in terms of age, education, race, and sex has remained relatively stable between 1985 and 2004. Furthermore, McPherson et al. (2001) noted that the mean heterogeneity of discussion networks is significantly less than the mean heterogeneity of the overall population, which reaffirms the idea that discussion networks are homophilous.

In political science, researchers often include demographic characteristics as control variables in their models of political behavior. But this practice does not capture the most theoretically interesting aspect of these characteristics – that race/ethnicity, age, religion, education, occupation, and gender provide contexts in which individuals think and talk about politics. Following Blau and McPherson, demographic characteristics can be viewed as dimensions of a social structure that defines social positions relative to one another. This departs from the conventional view that demographic characteristics are simply indicators of group interests or personal attributes, while accounting for the contextual analytical view that demography can have contextual effects.

In Blau space, great Blau distances indicate very low probabilities of contact. According to McPherson, "[t]his multidimensional Blau space at once organizes the social interactions among individuals, and structures the opportunities for the formation of social entities that are associated with individuals in that space." This organizing principle is the general rule for how individuals interact in the Blau space. The principle of homophily suggests that the more similar individuals are, the more likely they are to interact. In other words, the frequency of contact is inversely proportional to the Blau distance between any two individuals. When two individuals come into contact frequently, they are more likely to encounter or talk with each other. By encountering or talking with each other, they are more likely to transmit political attitudes than those who never or infrequently meet or talk.

Ideological space and distance

Ideological space (or "policy space" or "political choice space") is a latent, politically-defined space. Following the work of Melvin Hinich and his colleagues, ideological space is "a commonly-held simplification of the complex network of government policies and political issues." All voters are located on points on a k-dimensional ideological space. Based on previous research, there are $k = 2$ ideological dimensions where the first dimension is the well-known left-right ideological scale and the second dimension is a less explicit issue that changes from time to time, such as foreign policy or social issues.

According to the spatial theory of electoral politics, there is a relationship between the points in the ideological space and issues that become salient during election campaigns. In this ideological space, there are two groups of self-interested actors: voters and politicians. Voters are interested in the outcome of any given election, and they want to vote for the politician who is closest to their position in the space. This motivation is rooted in utility; the farther away a politician is from a voter's point in the space, the lower the utility associated with voting for that politician. Politicians are interested in winning the election and want to offer a "package" of policy statements, candidate characteristics, past record, and so on that would appeal to the greatest number of voters. Each politician's package defines his position in the ideological space. During a campaign, politicians and voters "interact" in this space, as politicians attempt to influence voters to vote for them and voters attempt to vote for politicians closest to their ideal issue positions.

The idea of ideological space from the spatial theory of voting can be very useful to the study of social influence. For the purposes of placing survey respondents on a common ideological space, the exact nature of the dimensions is not essential. What does matter is the spatial distance between voters in the space, which can be acquired simply by calculating the Euclidean distance between any two voters in the ideological space. In the spatial theory of voting, the distances between voters and candidates are of primary concern; in studying social influence, we are more concerned with the distances between voters because we are interested in understanding how the relationships among voters affect them as voters. In the spatial theory of voting, the distance between a given voter and a given politician impacts the voter's utility in voting for that candidate; in the context of social influence, the distance between two voters impacts the extent to which those voters will influence each other's political attitudes. Why might this be?

It is well-known that voters are not as informed as political scientists would like them to be (Converse, 1964). As in the spatial theory of voting, we assume here that voters are interested in election outcomes, which implies a certain level of political knowledge. But acquiring political knowledge is costly in terms of time and effort, and most people are not sufficiently interested in politics to sustain those costs. To overcome the information costs, individuals turn to sources that they trust to provide information that is already filtered and processed (Downs, 1957). Research has shown voters tend to seek out political experts (real or perceived) when they want political information and that these political experts tend to share the same ideological outlook. This suggests that voters are more affected by those who are closest to them in an ideological space and less influenced by those who are ideologically farther away. Thus, we can expect that any given Republican is more likely to be influenced by other Republicans than by Democrats, and that any given Democrat is more likely to be influenced by other Democrats than by Republicans. Furthermore, we can expect Republicans to visit websites and watch television programs that lean to the right, and we can expect Democrats to visit websites and watch television programs that lean to the left.

Geographical space and distance

In geographical studies, space refers to actual physical areas. But even geographical space need not be limited to geography. While paradoxical, this idea is borne out by how geography actually plays out in politics and social life.

While Huckfeldt used the census tract to approximate the set of unavoidable social interactions of everyday life, the census tract itself is meaningless. It is an artificial geographical boundary drawn for occasional census-taking purposes. Individuals inhabit this space in the sense that they are physically on it. But is this the only sense? For example, a "neighborhood" in the conventional sense has both geographical and social boundaries; it is defined physically by its buildings, parks, and sidewalks, but it is defined more meaningfully by its inhabitants, their relationships with one another, and their relationships with the physical space. While Huckfeldt stripped the definition of neighborhood so that it does not necessarily include a cohesive sense of community, he nonetheless included the day-to-day interactions among individuals who reside in the same geographical location. But these interactions are not necessarily constrained by physical proximity. Community norms and values also guide those interactions because geographical areas are where people grow up and live and necessarily acquire a sense of who they are in the context of fellow residents. There need not be a cohesive sense of community, but it is unrealistic to assume that individuals have absolutely no connection to where they live. First, individuals choose to move or stay in a geographical area for reasons that are social, such as job opportunities and school quality. When it comes to political science, these social qualities are the primary distinctions among geographical areas, rather than their size, elevation, and other physical characteristics. Second, geographical areas can include media markets, historical landmarks, and judicial circuits. These areas are nominally geographical but primarily functional. For these reasons, geographical spaces should be viewed as types of social spaces in political science research.

An election district is a type of geographical space. Individuals potentially interact with other individuals in the same election district, and individuals who live in adjacent election districts might also interact. But beyond that type of interaction, individuals may also have some idea of the norms and values and social realities of their election district and neighboring districts. Furthermore, when politicians and political groups reach out to voters, they tend to mobilize around geographical areas like election districts. For example, an environmental protection group might focus its organization, mobilization, and fundraising efforts in the areas surrounding a polluted lake but not much beyond it. In this situation, we expect more interaction among individuals in the areas surrounding the polluted lake because of this political effort, and less interaction between individuals in areas surrounding the lake and individuals in areas not surrounding the lake. In this way, geography has social meaning; geographical proximity becomes a kind of social proximity, and the geographical linkages become social linkages. In political science research, geographical relationships are often of this social nature, which is why it makes sense to subsume geographical space under the broader category of social space.

In contrast to Blau and ideological distances, geographical distance is a straightforward measure. Furthermore, with the increasing availability of geo-coded data, measuring geographical distance is more and more feasible for many types of spatial units.

Other spaces

Political scientists are generally interested in connections or ties among actors that are defined by some political or social phenomena. We can call the set of connections or ties defined by a particular political or social phenomenon a "social space" with dimensions corresponding to elements of that phenomenon. Subsets of this space are "spatial neighborhoods," which consist of points that are in some sense "close," such as degree of similarity or common membership. By viewing a set of politically or socially-based connections as a social space, we can represent these connections in a spatial weights matrix for spatial econometric analysis

While this chapter focused on Blau, ideological, and geographical spaces, the idea of social space is not limited to these three types. Previous studies have shown that social spaces can be based on *language similarity* (Dow et al., 1984), *trade or group membership* (Simmons and Elkins, 2004), *occupation and township* (Lin et al., 2006), *trade volume* (Beck et al., 2006), and *dyadic membership* (Beck et al., 2006). Social spaces can also be based on *historically shared ties* or *common acquaintance* (Beck et al., 2006). Furthermore, previous studies of social influence have focused on aggregate contexts and social networks that can also be considered social spaces. Aggregate contexts can be viewed as geographical spaces, since they tend to be census tracts, counties, states, and other geographically-bounded areas. Social networks can also be interpreted in the context of social space. We can call the set of all social networks a social space where spatial neighbors are those who mutual recognize each other as friends or discussants, which means that each neighborhood is a social network. The choice of which social space to study is a research consideration along the same lines as the unit of observation and choice of dependent variable.

ATTITUDES ARE CONTAGIOUS: A THEORY OF SOCIAL INFLUENCE

This section features a theoretical argument for how individuals in a social space might influence each other's political attitudes. First, I explain the concept of social influence and why social space is a useful tool for understanding social influence. Next, I explain the mechanisms of social influence. Finally, I identify testable hypotheses for gauging social influence in Blau, ideological, and geographical spaces.

Political attitudes are attitudes about political things, such as government spending and environmental protection. Following Crano and Prislin (2006), attitudes are "evaluative judgments" that integrate and summarize cognitive and affective reactions to an object. While most people, most of the time, do not think very much or very hard about politics, they do pick up bits and pieces of information and feelings about politics, and it is reasonable to suppose that they do so because politics is sometimes a topic of conversation with their families, friends, and colleagues. Previous research studies have shown that individuals do indeed rely on others for understanding politics. Individuals use their social networks as a shortcut for gathering political information (Huckfeldt, 2001; Huckfeldt and Sprague, 1995). Luskin et al. (2002) showed that the policy preferences of individuals can change quite a bit as a result of exposure to deliberation within small groups. Binder et al. (2009) found that having discussions with others affects individuals' attitudes toward stem cell research. Cho (2003) uncovered evidence of a diffusion effect in Asian-American campaign contribution networks. These and other studies show that social influence takes on a variety of forms, including network effects, small group effects, neighborhood effects, peer effects, and imitation.

Social influence as mutual influence

Social influence has been conventionally conceived as a one-directional effect of networks, small groups, neighborhoods, and other social groups on individuals. This view is based on the assumption that the effects of these groups are exogenous, or that individuals are influenced by these groups but do not, in turn, influence the groups to which they belong. Erbring and Young's (1979) work on the formal specification and substantive mechanisms of contextual effects shows that the assumption that social influence (as the effect of groups on individuals) is exogenous is both incorrect and unfortunate, for it underlies the inability of researchers to translate contextual effects into contextual explanations. To support the development of better contextual explanations, Erbring and Young argued that a contextual effect (*i.e.*, social influence) must be conceptualized as "a consequence of processes of interaction among individuals in a social network" – that is, social influence *is* mutual influence. Because this idea is important to this dissertation's treatment of social influence, I discuss Erbring and Young's argument in some detail here.

In their article, Erbring and Young first looked at the substantive mechanisms implied in conventional contextual models, and then looked at the structural specifications implied in conventional substantive interpretations of contextual effects.

In the first case, they found that the substantive mechanism implied in a typical contextual model is a kind of "social telepathy." In a typical contextual model, the individual-level dependent variable is the outcome of a combination of individual-level and group-level effects. The "contextual effect" is the effect of some group-level explanatory variable, which is usually represented by a group mean of a characteristic of interest. This effect can be interpreted as a direct effect of the group on the individual. In their example, Erbring and Young used individual academic achievement as the dependent variable y_{ij}, intellectual ability as the individual-level explanatory variable x_{ij}, and mean intellectual ability (of the *i*th individual's class *j*) as the group-level explanatory variable $\bar{\mathrm{x}}_{\cdot j}$:

$$y_{ij} = \mathrm{a} + \mathrm{b}_1 \mathrm{x}_{ij} + \mathrm{b}_2 \bar{\mathrm{x}}_{\cdot j} + \mathrm{e}_{ij}$$

According to Erbring and Young,

> [The equation above] implies the direct flow of effects from (each) student (i')'s ability to student (i)'s performance [where $\neq i'$], and similarly from student (i)'s ability to (each) student (i')'s performance *without*, in either case, letting this impact on performance be mediated by the student's own ability.

Noting that there may be "perverse situations" (*e.g.*, universal cheating) in which one student's intellectual ability directly affects another student's academic performance, Erbring and Young maintained that such a direct effect is generally implausible because there is no direct link between the two. Indeed, the presumed relationship between a group-level characteristic (such as intellectual ability) and individual-level outcome (such as academic performance) amounts to a kind of "social telepathy" that is unworthy of social science. "Yet 'action at a distance' is a well-known principle of magic, not of science which, on the contrary, is premised on the denial of that possibility and the search for intervening links," they wrote.

In the second case, Erbiring and Young found that two conventional substantive mechanisms that produce group effects – namely, *common fate* and *group norms* – actually implied structural specifications that did not correspond with the conventional contextual model. According to the common fate explanation, students in the same class share a "common fate" due to having the same teacher, sharing the same resources, or some other attribute of common class membership. But if this explanation is true, then it would make more sense to use one of these attributes as an explanatory variable rather than using mean intellectual ability as a proxy for it.

According to the group norms explanation, a group-level explanatory variable represents a kind of group "climate" that affects individual-level outcomes. According to Erbring and Young, however, this substantive explanation actually implies a structural specification without any group-level explanatory variables at all. Using individual academic achievement as the dependent variable y_{ij} and educational aspiration as the individual-level explanatory variable x_{ij}, Erbring and Young pointed out that, under the group norms explanation, the educational aspirations of one's peers affects one's *own* educational aspirations, but does *not* directly affect one's educational outcome. This mechanism has the following structural specification:

$$y_{ij} = \mathrm{a} + \mathrm{b}_1 \mathrm{x}_{ij} + \mathrm{e}_{i'j'} \qquad \text{where } i \neq i', j \neq j'$$

The problem with this structural specification is that the "contextual effect" due to the mutual influence on the explanatory variables cannot be estimated. Erbring and Young note:

> Unlike the intervening mechanisms considered previously, the present case – feedback among the "exogenous" variables – involves a genuine contextual process: interaction among individuals within a particular social structure. However, the

> effects of that process remain entirely hidden from view: the impact of social context on individual outcomes is confined to variables which are "exogenous" with respect to the outcome variable of interest . . . .

To remedy the problem of representing meaningful mechanisms of social influence, Erbring and Young proposed a class of "endogenous feedback models" that represent processes of mutual influence in a social context. These models represent social contagion processes, which are based on "reciprocal influence." In contrast to the conventional assumptions of the contextual model, an endogenous feedback model is based on the idea of "reciprocal influence":

> Individual behavior is assumed to be both passively responsive to the contextual cues provided by the behavior of significant others, and at the same time actively impinging upon the behavior of others sharing the same social environment.

A key characteristic of an endogenous feedback model is that the outcomes of interest are interdependent. In other words, individuals mutually influence each other through their dependent variables, resulting in interdependent individual outcomes within a social space. But before we look more closely at such an endogenous feedback model, it is necessary to discuss in further detail the mechanisms of mutual influence.

Mechanisms of mutual influence

There are several explanations for how social influence (as mutual influence) works; these explanations are generally based on the idea that social interaction, however defined, has consequences for individual behaviors and attitudes. According to the social conformity explanation, for example, individuals generally do not want to come into conflict with others because of social or psychological costs. That means individuals will tend to agree – or at not least not disagree – with the views of others. According to the group identity (or reference group) explanation, individuals have certain social loyalties, and these loyalties color their perception of political information and issues. According to the impersonal influence theory, individuals develop their perceptions of what others think from information they get from the media and use this perception to form or change their political attitudes and behaviors (see Mutz, 1992). According to the information cascade theory, whether an individual decides to act (or not act) depends on whether the number or proportion of other individuals acting (or not acting) exceeds a certain threshold (Granovetter, 1978). This suggests, for example, that an individual might adopt a pro-environmental attitude if enough other individuals do so as well. Despite their different focuses, all of these explanations of social influence suggest that when individuals form or change their attitudes toward politics, they take into account other individuals' attitudes as well. All these various theories suggest that one's susceptibility to social influence is *strategic*.

According to Franzese and Hays (2007a) and Brueckner (2003), strategic interdependence arises whenever some unit(s)'s actions affect the marginal utilities of the alternative actions for some other unit(s). Their general theoretical model of strategic interdependence can be easily adapted to social influence.

Suppose there are two individuals (which we denote as individual *i* and individual *j*), and suppose they derive indirect utilities (which we denote as U_i and U_j) from their respective political attitudes. Because of externalities resulting from social interaction, individual *i*'s utility U_i depends both on his own attitude a_i and individual *j*'s attitude a_j so that

$$U_i \equiv U_i(a_i, a_j)$$
$$U_j \equiv U_j(a_j, a_i)$$

What this means is that when an individual *i* takes on a political attitude a_i, his choice depends on the political attitude a_j of individual *j* and vice versa. The strategic interdependence between individuals *i* and *j* can be expressed with two best-response functions such that individual *i*'s optimal political attitude a_i' is a function of individual *j*'s chosen political attitude a_j and vice versa:

$$a_i' \equiv Argmax_{a_i} U_i(a_i, a_j) \equiv R_i(a_j)$$

$$a_j' \equiv Argmax_{a_j} U_j(a_j, a_i) \equiv R_j(a_i)$$

The slopes of the functions provide information on the relationships between individuals *i* and *j*. If the slopes of these best-response functions for individual *i* is positive, then the political attitude of individual *j* causes individual *i* to move in the same direction. In this case, Franzese and Hays (2007a) would refer to their attitudes as "strategic complements." On the other hand, if the slope for individual *i* is negative, then the political attitudes of *i* and *j* move in the opposite direction. In this case, the attitudes would be "strategic substitutes."

When it comes to social influence in politics, we should generally expect political attitudes among mutually influential individuals to be "strategic complements." Psychological studies of attitude formation and change suggest that individuals tend to conform to rather than contradict other individuals in their social environment.

Social influence can lead individuals to form new attitudes or change existing ones. Mechanisms underlying attitude *formation* can be categorized by the degree of conscious awareness (Crano and Prislin, 2006). An example of a "below-conscious" mechanism is mere exposure. According to the theory of mere exposure, the frequency of exposure to something tends to increase an individual's affect toward that thing. In the context of political attitudes, the idea of mere exposure suggests that, by simply being around other people, individuals are thereby exposed to their attitudes and gradually come to adopt those same attitudes themselves. This also suggests that others' attitudes are influential not only because of their credibility or content, but because of the nature of sociability itself. People want to be liked by other people, and accordingly, they are inclined to conform in opinion and action. This inclination may be conscious or unconscious. In their study of automaticity, Bargh and Chartrand (1999) argue that individual attitudes and behaviors do not originate from within the individual but are largely caused by the automatic, unconscious processing of environmental factors. This automaticity means that the internal weighing of pros and cons is not necessarily the primary cause of an individual's position on political issues. In fact, these evaluations may be more automatic than commonly realized (Bargh and Chartrand, 1999). Accordingly, many political attitudes may not be conscious judgments but automatic evaluations. Since most people do not think about politics most of the time, the concepts of mere exposure and automaticity are consistent with empirical findings that many political attitudes are lightly held and poorly thought out.

Social groups can bring about attitude *change* by enhancing an individual's receptivity to arguments and by limiting the pool of arguments. According to Druckman and Lupia's (2000) review of preference formation, persuasion can be divided into three components: recipient effects, message effects, and source effects. Recipient effects pertain to the characteristics (such as political sophistication) of individuals that affect how receptive they are to persuasive attempts. Message effects refer to the characteristics of persuasive appeals, such as their content and tone. Finally, source effects refer to the characteristics of the source of persuasive appeals, such as a speaker's trustworthiness, popularity, insider status, accuracy and objectivity, and ideology. Social influence can be seen as a type of source effect. Because persuasion depends, in part, on the personal attributes of those attempting to persuade (Lupia 2002), would-be persuaders need to be (or appear to be) trustworthy and knowledgeable before the actual argument is even heard at all.

The literature regarding attitude formation and change points to the importance of individual characteristics in transmitting and assimilating attitudes. By combining this idea with Erbring and Young's idea that social influence must be mutual influence and with our adaptation of Franzese and Hay's general theory of strategic interdependence, we come to a working theory of social influence: Social influence is a process that leads to the formation of or change in an individual attitude (or behavior) within a social space. It takes place when a group of individuals in a social space mutually influence each other's attitudes (or behaviors) such that each individual finds it advantageous to adjust their attitudes (or behaviors) in the direction of other individuals in the group. This theory underlies processes such as social contagion, facilitation, competition, conformity, diffusion, emulation, learning, and coercion.

MODELING SOCIAL INFLUENCE

Based on our working theory, a formal specification of social influence must have certain features that are currently missing from conventional specifications (*e.g.*, contextual models). Mutual influence among individuals must refer to mutual influence among the individuals' outcomes of interest, which, according to Erbring and Young (1979), is called "endogenous feedback." This means that the dependent variable outcomes are interdependent, while allowing for a number of exogenous independent variables. Erbring and Young (1979) provide an example of such an "endogenous feedback model":

$$y_{ij} = \alpha \left(\sum_{j'=1}^{m} \sum_{i'=1}^{n_{j'}} w_{(ij)(i'j')} y_{i'j'} \right) + \beta_0 + \beta_1 x_{ij} + u_{ij}$$

where $i = 1, \ldots, n_j; j = 1, \ldots, m$

In this endogenous feedback model, the dependent variable y_{ij} is an outcome associated with individual i, who is associated with group j. y_{ij} is a function of an individual's personal characteristic xij *as well as* the outcomes of the other individuals in group *j*. The parameter α is the "feedback coefficient" that represents the effect of mutual interaction among the individuals in a group, while the parameter β_1 represents the effect of the individual-level characteristic x_{ij}. The term $w_{(ij)(i'j')}$ represents the "extent of interaction between any two students; it is equal to 0 if they do not interact and if $i = i'$and $j = j'$. The error term u_{ij} is assumed to be independently and normally distributed.

According to Erbring and Young, the endogenous feedback model has two important features:

> First, it includes terms in $y_{i'j}$ in the equation for y_{ij}, where (*i'*) represents the relevant peers with whom (*i*) interacts; this has important implications for the error term, u_{ij} . . . . Second, the model incorporates explicit assumptions for each (*i*) about who the relevant peers (*i'*) are; thus the social structure through which the particular effects are mediated becomes an integral part of the specification of the model, as suggested by the $w_{ii'}$ terms (and their associated coefficient, α).

In other words, the endogenous feedback model features the dependent variable on both sides of the equal sign and the social structure that provides the opportunity for social influence. It turns out that Erbring and Young's endogenous feedback model is actually a version of the spatial lag model, which is a widely-used model in the field of spatial econometrics.

INTRODUCING SPATIAL ECONOMETRICS

Rooted in geography, early work in spatial econometrics was limited to the statistical analysis of regional and urban modeling using geographical data. It was not until the mid-2000s that interest in spatial econometrics went beyond the bounds of regional science. In 2006 alone, Anselin (2006) established spatial econometrics in the methodological toolbox of applied econometrics; Lin et al. (2006) expanded the idea of spatial dependence to mutual influence within townships and occupations in a study of national identity in Taiwan; and

Beck et al. (2006) argued that spatial econometric models should not be limited to geography and can accommodate political economy notions of distance. Two years later, Ward and Gleditsch (2008) released *Spatial Regression Models*, an introduction to spatial econometrics written for social scientists. Understanding social influence within the spatial econometric framework, as this dissertation does, is a logical step forward in advancing the use of spatial econometrics in political science.

Spatial econometrics is a set of tools for working with data with spatial effects, which include spatial dependence and spatial heterogeneity. *Spatial dependence* is "the existence of a functional relationship between what happens at one point in space and what happens elsewhere" (Anselin, 1988). *Spatial heterogeneity* refers to a situation in which there is spatial dependence but the relationship between values and locations is unstable – that is, "functional forms and parameters vary with location and are not homogenous throughout the data set" (Anselin, 1988). This dissertation is concerned with the first type of spatial effect – *i.e.*, spatial dependence - and how it can be expanded to include the concept of social influence.

The Concept of Spatial Dependence

Also called "spatial interdependence," spatial dependence is a fundamental concept in spatial econometrics. According to Anselin (1988), spatial dependence is based on the idea that relative distance has an effect on observed quantities of interest. This effect tends to arise from two main sources: measurement error and spatial interaction. Measurement error can give rise to spatial dependence when data are organized in space or in space and time. Anselin (1988) elaborates:

> In a more general sense, any situation where data are structured subject to a measure of location or distance, in any space, may be considered. For example, measures

derived for clusters of industrial activities characterized by a sectoral profile (*i.e.*, a point in multi-dimensional profile space), or information on nations or [interest] groups identified by their positions in policy space would satisfy this requirement.

The second and more substantively interesting source of spatial dependence is spatial interaction. Spatial interaction follows from "the importance of space as an element in structuring explanations of human behavior" (Anselin, 1988). Because human beings live in space and time, their relative location and distance matter in a variety of ways (*e.g.,* diffusion processes.) and as a result, what is observed as one point in a space is, in part, determined by what happens elsewhere in the space. Formally, this can be expressed as follows:

$$y_i = f(y_1, \ldots, y_n) \quad i \in S$$

where S is the set containing all spatial units of observation and f is the function relating each observation of y for one spatial unit to the values of y for the other spatial units. As Anselin (1988) points out, the function f is unidentified, but by imposing a form for the spatial process it represents, one may estimate and test empirically some characteristics of the spatial dependence at hand. Specifying the form of spatial dependence, incorporating the form into a formal model specification, and using the model for data analysis are the primary tasks of spatial econometrics.

The Spatial Weights Matrix

The first step of specifying the form of spatial dependence involves specifying the form of the spatial weights matrix (also called a "connectivity matrix"), which is denoted by **W**. **W** is an n by n spatial weights matrix that represents the connectivities among spatial units, such as countries, cities, and individuals. A "connectivity" is a researcher-defined relationship between two spatial units; for example, two countries can be considered connected if they share a common border, if they are trading partners, or if they share a common language. This connectivity is assumed to be known to the researcher, which means that the specific spatial weights matrix to be used in a spatial econometric analysis is not estimated. While this is a strong assumption, Beck et al. (2006) point out that "[I]t is no stronger than the typical implicit assumption that all connectivities are zero, that is, all observations are spatially independent."

The connectivity between individuals i and j is denoted by the element $w_{ij}*$. The simplest type of connectivity is binary membership, where

$$
\begin{aligned}
&w_{ij}* = 1 \quad \text{if individuals } i \text{ and } j \text{ belong to the same group and } i \neq j \\
&w_{ij}* = 0 \quad \text{otherwise}
\end{aligned}
$$

For example, if individual 1 and individual 2 are members of the same civic organization and individual 3 is a member of a different civic organization, then

$$
\begin{aligned}
&w_{11}* = 0 \\
&w_{12}* = 1 \\
&w_{13}* = 0 \\
&w_{21}* = 1 \\
&w_{22}* = 0 \\
&w_{23}* = 0 \\
&w_{31}* = 0 \\
&w_{32}* = 0 \\
&w_{33}* = 0
\end{aligned}
$$

These weights can then be included in the following spatial weights matrix:

$$
\mathbf{W}_{n \times n}* = \begin{bmatrix} 0 & 1 & 0 & \cdots & w_{1n}* \\ 1 & 0 & 0 & \cdots & w_{2n}* \\ 0 & 0 & 0 & \cdots & w_{3n}* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1}* & w_{n2}* & w_{n3}* & \cdots & 0 \end{bmatrix} \tag{2.1}
$$

Note that the diagonal elements of (2.1) are all equal to 0, which is conventional in spatial econometrics. When individuals i and j belong to the same group (*i.e.*, $w_{ij}*$ is a nonzero number), they are said to be "neighbors" and, by convention, no individual is a neighbor to himself. Each group of individuals is called a "neighborhood."

Other popular types of spatial weights matrices are those based on inverse distance, inverse distance squared, distance band, and k-nearest neighbor. In spatial matrices based on inverse distance or inverse distance squared, the elements are functions of the distance (*e.g.*, miles) between individuals i and j. In a distance band matrix, $w_{ij}*$ is non-zero if the distance between individuals i and j meets some threshold, such as 500 or greater miles. If the distance between individuals i and j does meet the threshold, then $w_{ij}*$ can be assigned a value of 1 or some function of the distance. In a k-nearest neighbor spatial weights matrix, the element $w_{ij}*$ is non-zero if individual j is one of the k nearest neighbors to individual i, where k is a positive integer and "nearest" is based on some type of distance.

Conventionally, the spatial weights matrix is row-standardized. Each element $w_{ij}*$ is divided by its corresponding row sum so that the elements of each row sum to one. Throughout this dissertation, $\mathbf{W}*$ will denote an unstandardized spatial weights matrix, while $\mathbf{W}$ will denote a row-standardized spatial weights matrix.

The spatial weights matrix is essential to spatial analyses. After specifying spatial weights matrix, one can then measure and assess the existence of spatial dependence using spatial autocorrelation and spatial regression analysis.

SPATIAL REGRESSION MODELS

Spatial regression models allow researchers to assess whether there is still spatial dependence in a variable of interest (detected in tests of spatial autocorrelation) after controlling for relevant variables. For example, a researcher might use a spatial regression model to see whether there is spatial dependence in vote choice across counties while controlling for party affiliation and social-economic status.

While there are many types of spatial regression models, the spatial lag model is most consistent with the idea that social influence is a diffusion process. Therefore, it is the workhorse model of this dissertation. Other commonly-used models include the first-order spatial autoregressive model, spatial-x model, spatial error model, and spatial simultaneous autoregressive model, which will also be briefly discussed below.

The Spatial Lag Model

Also called the "mixed regressive autoregressive model" or "spatial autoregressive model," the spatial lag model is the primary model used in this dissertation. The spatial lag model takes the following form:

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \varepsilon &\sim N\left(0, \sigma^2\mathbf{I}\right) \end{aligned} \tag{2.4}$$

In this model, $\mathbf{y}$ is the *n* by 1 vector of observations of the dependent variable. $\boldsymbol{\beta}$ is a *k* by 1 vector of parameters associated with independent variables $\mathbf{X}$, which is an *n* by *k* matrix. The error term $\boldsymbol{\varepsilon}$ satisfies the usual Gauss Marcov assumptions, and $\mathbf{I}$ is a *n* by *n* identity matrix. Note that this is essentially the same model as Erbring and Young's endogenous feedback model if we assume that $\mathbf{W}$ is based on binary membership.

The spatial lag model is characterized by the spatial lag term $\mathbf{Wy}$. The spatial lag $\mathbf{Wy}$ is an *n* by 1 vector where each element is the lagged values of the dependent variable. Depending on how $\mathbf{W}$ is defined, the spatial lag can be interpreted as the unweighted mean, weighted mean, or sum of the *y* values of respondent *i*'s neighbors. The spatial autoregressive parameter ρ measures the effect of the spatial lag. If $\rho = 0$, then there is no spatial dependence, and the spatial lag model reduces to the classical linear regression model.

On the other hand, if the spatial autoregressive parameter ρ is statistically different from zero, there would be evidence of a neighborhood effect on political attitudes.

The First-Order Spatial Autoregressive Model

While the first-order spatial autoregressive (FAR) model is seldom used in applied work, it is useful for exploring theoretical ideas and concepts (see LeSage, 1998). The FAR model is given by:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \sigma^2\mathbf{I}_n\right)$$

In the FAR model, the variable vector $\mathbf{y}$ is expressed in deviations from the means. $\mathbf{Wy}$ is the spatial lag of $\mathbf{y}$. The spatial autoregressive parameter ρ is the effect of $\mathbf{Wy}$ on $\mathbf{y}$ – that is, the extent to which variations in y are explained by neighboring y values. The error term $\boldsymbol{\varepsilon}$ satisfies the usual Gauss Marcov assumptions, and $\mathbf{I}$ is a n by n identity matrix.

The Spatial-X Model

In the spatial-x model the dependent variable $\mathbf{y}$ is regressed on two types of explanatory variables: the spatially-weighted explanatory variables $\mathbf{WZ}$ and the conventional explanatory variables $\mathbf{X}$.

The spatial-x model takes the form

$$\mathbf{y} = \gamma\mathbf{W}\mathbf{Z} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} \sim N\left(0, \sigma^2\mathbf{I}\right)$$

where $\mathbf{y}$ is an n by 1 dependent variable vector, $\mathbf{Z}$ is spatially-relevant explanatory variable vector, $\mathbf{WZ}$ is the n by 1 spatial-x term, and $\mathbf{X}$ is the n by k independent variable matrix. The spatial parameter γ is the effect of the spatial-x term $\mathbf{WZ}$, while $\boldsymbol{\beta}$ is a k by 1 vector of parameters associated with independent variables **X.** As usual, the error term ε is assumed to satisfy the usual Gauss Marcov assumptions, and **I** is a *n* by *n* identity matrix.

Note that in contrast to the spatial lag model and FAR model, the spatial term in the spatial-x model is exogenous. This means that the effect of a social content is unidirectional; the individual is influenced by the social context but does not influence the social context.

The Spatial Error Model

The *spatial error model* takes the following form:

$$
\begin{aligned}
Y &= \mathbf{X}\beta + \varepsilon \\
\varepsilon &= \lambda \mathbf{W} \varepsilon + \mu \\
\mu &\sim N(0, \sigma^2 \mathbf{I})
\end{aligned}
$$

This model can account for the possibility that a given social context constitutes a spatially correlated omitted variable, causing the errors of a model to be spatially correlated. β is a *k* by 1 vector of parameters associated with independent variables **X**, which is an *n* by *k* matrix. The error term ε has been decomposed into two parts: the spatial component $\lambda\mathbf{W}\varepsilon$ and the component μ, which satisfies the usual Gauss Marcov assumptions. **I** is a *n* by *n* identity matrix, and the errors μ are distributed normally with a constant variance. The term $\lambda\mathbf{W}\varepsilon$ accounts for the average error associated with each social context. The spatial error parameter λ indicates the extent to which the errors are correlated for nearby observations. If there is no spatial correlation between observations *i* and *j* $(i \neq j)$, then λ will equal 0, and the model will reduce to the classical linear regression model.

The Spatial Simultaneous Autoregressive Model

The spatial simultaneous autoregressive model takes the following form:

$$Y = \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X}\beta + \varepsilon$$
$$\varepsilon = \lambda \mathbf{W}_2 \varepsilon + \mu$$
$$\mu \sim N\left(0, \sigma^2 \mathbf{I}\right)$$

This model combines the spatial lag model with the spatial error model. The advantage of the spatial simultaneous autoregressive model is that it allows for the simultaneous estimation of the spatial autoregressive parameter ρ and the spatial error parameter λ, which allows for the evaluation of two different contexts (when $\mathbf{W}_1 \neq \mathbf{W}_2$) with only one model.

COMPETING MODELS OF SOCIAL INFLUENCE

To study social influence, researchers require data on groups and on individuals. The analysis of data collected at different levels falls under the broad category of multilevel analysis. The most popular types of multilevel analysis for studying social influence are the analysis of variance (ANOVA), contextual analysis, and social network analysis.

Analysis of Variance

ANOVA is a method for finding out whether different groups have different characteristics, values, behaviors, or some other quality of interest. For example, we can use ANOVA to find out whether Protestants, Catholics, and Unitarians have different levels of political participation and whether those differences are statistically significant and meaningful. To do this, we can use the following fixed one-way ANOVA model (assuming that we have data on all three groups):

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

Each y_{ij} represents an observation of an individual's level of political participation. The term μ represents the overall mean level of political participation, α_j represents the mean level of political participation for the jth group, and ε_{ij} represents the individual-level random error.

According to the ANOVA framework, if different groups have different means (under the conventional assumption that the variance is constant), the between-group variance should be greater than the within-group variance and we use the *F*-test to see if the data meets this expectation. If it turns out that the group differences are statistically significant, we can then conclude that being a Protestant or a Catholic or a Unitarian has an effect (denoted by α_j) on an individual's level of political participation and that this is some evidence of a group-level effect, or social influence.

ANOVA, or dummy variable regression, can be viewed as a special case of spatial regression. This will be shown in a later chapter, which features a detailed discussion of the differences and similarities between linear regression and spatial regression. For now, suffice it to say that the ANOVA approach is limited to finding out whether different groups are merely different; it says nothing about *why* those groups are different. This is why researchers often turn to contextual analysis to study social influence.

Contextual Analysis

Contextual analysis is a popular method for finding out whether particular characteristics of a social, political, economic, or some other type of context affects how individuals think or behave. By using the ANOVA framework, for example, researchers can find out whether Protestants, Catholics, and Unitarians have different levels of political participation, whereas a contextual analysis can tell researchers whether an individual has a higher level of political participation due to some specific *characteristic* of Protestants, Catholics, or Unitarians. For example, a Protestant might have greater political participation when there is a large percentage of Protestants living in the same county, or a Catholic might have greater political participation when the average level of political knowledge among fellow Catholics is higher. Researchers can use a contextual model to find out if the effect of a higher percentage of Protestant county-dwellers or the effect of average Catholic political knowledge is a statistically significant factor in an individual's level of political participation. A contextual model for such a study might look something like this:

$$y_{ij} = \alpha_0 + \alpha_1 x_{ij} + \alpha_2 \bar{x}_j + \varepsilon_{ij} \tag{2.5}$$

Each y_{ij} represents an observation of an individual's level of political participation. In the model, α_0 is the overall intercept, α_1 is the effect of the individual-level independent variable of political knowledge x_{ij}, α_2 is the effect of the religious group-level independent variable of political knowledge $\bar{x}_j$, and ε_{ij} represents the individual-level random error. If the group-level effect α_2 is statistically significant, we can conclude that the level of political knowledge of an individual's religious group has an effect on the individual's level of political participation and that this is evidence of a group-level influence, or social influence. In contrast to the ANOVA findings, we can argue that political participation is subject to social influence because of the need for political knowledge, which religious groups might provide. In this manner, contextual analysis can help determine *why* groups matter.

Contextual models are special cases of the two-level hierarchical linear model, which Lin et al. (2006) have shown to be a special case of the spatial lag model. The proceeding discussion follows Lin et al.'s (2006) demonstration and uses it to show how spatial lag models can provide all the information that a contextual model can and more, thereby showing that spatial lag models are more than adequate for studying social influence.

Consider m groups, and each group comprises n_k members (k = 1, 2, … , m). Let $\mathbf{y}_k$ be the n_k observations for the kth group, $\mathbf{X}_k$ be the corresponding n_k by k matrix of independent variables, $\boldsymbol{\beta}$ be a k by 1 vector of coefficients, and $\boldsymbol{\varepsilon}_k$ be the n_k by 1 vector of independent, normally-distributed disturbances. Let ρ denote the spatial autoregressive parameter. Using the identity $\mathbf{Wy} \approx b_1\mathbf{D}_1 + b_2\mathbf{D}_2 + ... + b_m\mathbf{D}_m = \bar{\mathbf{y}}$, the spatial lag model

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \text{where } |\rho| < 1$$

is asymptotically equivalent to

$$\mathbf{y} \approx \rho\bar{\mathbf{y}} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

By taking the average across each neighborhood, we get

$$\bar{\mathbf{y}} \approx \upsilon\bar{\mathbf{y}} + \bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\boldsymbol{\varepsilon}}$$

or

$$\begin{bmatrix} \bar{\mathbf{y}}_1 \\ \bar{\mathbf{y}}_2 \\ \vdots \\ \bar{\mathbf{y}}_m \end{bmatrix} \approx \varrho \begin{bmatrix} \bar{\mathbf{y}}_1 \\ \bar{\mathbf{y}}_2 \\ \vdots \\ \bar{\mathbf{y}}_m \end{bmatrix} + \begin{bmatrix} \mathbf{1} & \bar{\mathbf{x}}_{11} & \dots & \bar{\mathbf{x}}_{k1} \\ \mathbf{1} & \bar{\mathbf{x}}_{12} & \dots & \bar{\mathbf{x}}_{k2} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{1} & \bar{\mathbf{x}}_{1m} & \dots & \bar{\mathbf{x}}_{km} \end{bmatrix} + \begin{bmatrix} \bar{\boldsymbol{\varepsilon}}_1 \\ \bar{\boldsymbol{\varepsilon}}_1 \\ \vdots \\ \bar{\boldsymbol{\varepsilon}}_m \end{bmatrix}$$

which can be rearranged into the following form:

$$\bar{\mathbf{y}} \approx \frac{1}{1-\rho}\left(\bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\boldsymbol{\varepsilon}}\right)$$

By substituting the above expression into $\mathbf{y} \approx \rho\bar{\mathbf{y}} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, we get

$$\mathbf{y} \approx \rho\left[\frac{1}{1-\rho}\left(\bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\boldsymbol{\varepsilon}}\right)\right] + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\mathbf{y} \approx \frac{\rho}{1-\rho}\left(\bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\boldsymbol{\varepsilon}}\right) + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

This is a two-level hierarchical linear model with fixed effects from neighborhood-level variables and a random intercept. If we assume a constant intercept $(\bar{\boldsymbol{\varepsilon}} = 0)$, it becomes a contextual model with fixed effects and no interactions:

$$\mathbf{y} \approx \frac{\rho}{1-\rho} \cdot \bar{\mathbf{X}}\boldsymbol{\beta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Note that the more familiar form of the contextual model is

$$\mathbf{y} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1\bar{\mathbf{X}} + \boldsymbol{\alpha}_2\mathbf{X} + \boldsymbol{\varepsilon}$$

The purpose of estimating a contextual model is to parse and compare individual and group-level effects – in other words, estimating the vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$. At first glance, the contextual model appears more desirable for these purposes than the spatial lag model. First, there are no *a priori* restrictions on parameter vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$. Second, there are minimal assumptions involved, since an individual observation is simply a linear combination of individual-level and group-level variables, and the error term follows the usual classical assumptions. Finally, there is no need to impose a structure on the relationships among individuals using the spatial weights matrix. Unfortunately, these "desirable" characteristics have led to a host of problems for the study of social influence.

First, empirical evidence suggests that there *is* a relationship between parameter vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$. Individual-level and group-level independent variables are usually collinear (see Blalock, 1984; Iverson, 1991). This collinearity problem implies the usual set of issues for significance testing. In the spatial lag model, the coefficient vector for $\overline{\mathbf{X}}$ is clearly constrained as a multiple of the coefficient vector of $\mathbf{X}$; the multiplier $\frac{\rho}{1-\rho}$ represents the relationship between the effects of the individual-level and group-level independent variables. Note that there are no contextual effects when $\rho = 0$, the individual-level effects are the same as the contextual effects when $\rho = 0.5$, and the contextual effects become infinitely large as ρ approaches 1. Furthermore, the parameter ρ represents the size of the group influence itself and can be subjected to significance testing. In contrast, the absence of ρ in the contextual model makes it difficult to compare the size of individual-level versus group-level effects, since the effects are estimated at different levels of measurement (using classical methods).

Second, the minimal assumptions required for estimating a contextual model come at a great cost: There is no explicit representation of the process of group influence.

Consequently, researchers debate what the processes of group influence might be, rather than subject these theories to empirical evaluation. The minimalism of the contextual model prohibits them from doing otherwise. In the spatial lag model, the presence of the **y** variable on both sides of the equal sign represents an endogenous process of group influence. An individual observation is influenced by other individual observations *and* influences other individual observations in turn. Furthermore, for individual i, a change in an individual-level independent variable affects y_i and, because of the endogeneity, also affects individual i's neighbors' y values as well. This process of mutual influence is why Lin et al. (2006) described the spatial lag model as "a dynamic spatial contagion process in equilibrium."

Finally, contextual models do not require researchers to specify group structures. While researchers often present theories of how individuals fit into groups, contextual models do not allow these social relationships to be represented. Consequently, empirical results derived from contextual models are too general to support or reject specific theories of group influence. In contrast, using a spatial lag model means having the capability to assess different group structures by specifying an appropriate spatial weights matrix, estimating the model, and testing the significance of the parameter ρ. An often-cited weakness of spatial regression models is that different spatial weights matrices produce different results; in the case of social influence, however, this can be a source of strength because it allows researchers to compare the results of different assumptions regarding social structures.

Social Network Analysis

In network analyses in political science, researchers often rely on social network data. This type of data is gathered using specially-designed surveys that are conducted to assess whether and to what extent discussing politics with social network members affects an individual's political attitudes and behaviors. Social network data includes information gathered on survey respondents as well as information gathered on individuals that survey respondents have identified as people with whom they discuss politics and/or important matters (often called "network members" or "discussants").

To assess whether and to what extent social network members affect an individual's political attitudes and behaviors, researchers use some form of the following model:

$$\mathbf{y} = \overline{\mathbf{Z}}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.6}$$

where $\mathbf{y}$ is an N by 1 dependent variable vector, $\overline{\mathbf{Z}}$ is an N by l matrix of network measures, and X is the N by k independent variable matrix. For clearer and simpler exposition, we assume that $l = 1$, which means that there is only one network measure. The scalar α is the parameter associated with the network measure, while $\boldsymbol{\beta}$ is a k by 1 vector of parameters associated with individual-level independent variables **X.** As usual, the error term ε is assumed to satisfy the usual Gauss Marcov assumptions. Note that N is the number of survey respondents and that $\mathbf{Z}$ is based on data acquired from the ith respondent's p network members (where $p \geq 2$ and p is assumed to be the same for all N respondents).

For assessing network effects, the parameter of interest is α, which is the effect of $\mathbf{Z}$ on $\mathbf{y}$. The network measure represented by $\boldsymbol{Z}$ is the measure of some characteristic of a respondent's social network, such as the mean level of political expertise, percentage of Republicans (or Democrats), or frequency of discussion. Such a characteristic is assumed to be exogenous to the survey respondents, and if it has a statistically significant effect, this is taken to be evidence of network influence on an individual attitude or behavior of interest.

An illustrative example of the social networks approach in political science is McClurg's (2006) study of social influence on political participation. McClurg examined three aspects of social networks: the level of political disagreement, the supply of political expertise, and the frequency of discussions about politics. Measures of these variables were obtained for each survey respondent's social network. To gauge the magnitude of these effects on individual political participation, McClurg included these network variables in his negative binomial regression model and found that they were statistically significant variables. His findings showed that higher levels of network agreement, more frequent discussions with network members, and higher levels of political knowledge in the network led to higher counts of political activities for individuals.

It turns out that network models such as (2.6) are special cases of spatial-x models, which we will see in the following discussion.

Suppose there is social network data on N survey respondents who each have p discussants. Assuming that no respondent shares a discussant, this means that there are $N + N * p = n$ total individuals in the survey. Suppose we are interested in a dependent variable y as a function of a network characteristic Z and k individual-level variables $\mathbf{X}$. Then the dependent variable y can be modeled using the following spatial-x model:

$$\mathbf{y}^* = \gamma\mathbf{W}\mathbf{Z} + \mathbf{X}\boldsymbol{\beta} + \mathbf{v}$$

where $\mathbf{y^*}$ is an n by 1 dependent variable vector, $\mathbf{X}$ is the n by k independent variable matrix, $\boldsymbol{\beta}$ is a k by 1 vector of parameters associated with independent variables $\mathbf{X}$, and the error term $\mathbf{v}$ is assumed to satisfy the usual Gauss Marcov assumptions.

Like other spatial regression models, the spatial weights matrix $\mathbf{W}$ is an integral part of the spatial-x model. The n by n matrix $\mathbf{W}$ is the row-standardized version of $\mathbf{W}^*$, which is defined as follows:

$w^*_{ij} = 0$ if individuals i and j are not in the same social network or $i = j$

$w^*_{ij} = 1$ if individuals i and j are in the same social network

Thus, $\mathbf{W}^*$ represents the social network structure of all n respondents. Since no individual can belong to more than one social network, $\mathbf{W}^*$ can be represented as a block-diagonal matrix with N disjoint neighborhoods. Following convention, we row-standardize $\mathbf{W}^*$ by dividing each element by the row sum – *i.e.*, p – to obtain $\mathbf{W}$.

Post-multiplying $\mathbf{W}$ by $\mathbf{Z}$ yields the spatial-x term $\mathbf{WZ}$. Because $\mathbf{W}$ is row-standardized, each ith element of the spatial-x term $\mathbf{WZ}$ represents the mean of Z for the ith individual's network members, excluding the value of Z of the ith individual. The parameter γ is the effect of $\mathbf{WZ}$ on $\mathbf{y}$; if γ is positive and statistically significant, then an individual-level attitude or behavior is subject to network influence because of a network characteristic represented by $\mathbf{Z}$.

Using the identity $\mathbf{WZ} = \overline{\mathbf{Z}}$, we can rewrite the spatial-x model as follows:

$$\mathbf{y}^* = \gamma\overline{\mathbf{Z}} + \mathbf{X}\boldsymbol{\beta} + \mathbf{v} \tag{2.7}$$

It is clear that (2.7) is a network model very similar to (2.6), since $\mathbf{y}^*$ is a function of $\overline{\mathbf{Z}}$, a vector of the means of a network characteristic Z, and of some individual-level variables $\mathbf{X}$.

It is important to note that there are several important differences between models (2.6) and (2.7). First, the sample sizes are different but related. Whereas (2.6) is based on N observations, (2.7) is based on $N + N * p = n$ observations. In other words, (2.6) is based on a subset of (2.7). Second, (2.6) contains observations of y and $\mathbf{X}$ only for the survey respondents, not for the discussants. In contrast, (2.7) contains observations of y and $\mathbf{X}$ for both survey respondents and discussants. Third, while the parameters γ and α are both network effects, γ is the network effect on all n individuals, while α is the network effect only for the N survey respondents.

In fact, model (2.6) can be viewed as a truncated version of model (2.7). The observations in (2.6) are selected based on whether the individuals are survey respondents and omitted when they are discussants. Equating (2.6) and (2.7) must therefore depend on some assumptions regarding their relationship.

Since the observations in (2.6) are selected on whether the individuals are survey respondents and omitted when they are discussants, we can represent this selection formally:

$$\begin{aligned} s_i{*} &= \mathbf{x}_{0i}'\boldsymbol{\alpha}_0 + u_i \\ s_i &= 0 \quad \text{if } s_i{*} \leq 0 \\ s_i &= 1 \quad \text{if } s_i{*} > 0 \end{aligned} \tag{2.8}$$

where $s_i = 0$ means that an individual *i* is a discussant and $s_i = 1$ means that an individual *i* is a survey respondent. The binary variable s_i is the realization of an unobserved variable $s_i{*}$, which is a function of some unknown independent variables $\mathbf{x}_0$ and is assumed to have a normally distributed, independent error term *u* with a mean of zero and constant variance. The observed variable *y* of model (2.6) is given by

$$\begin{aligned} y_i{*} &= \bar{z}_i\boldsymbol{\alpha} + \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i \\ y_i &= y_i{*} \quad \text{if } s_i = 1 \\ y_i & \text{ is unobserved if } s_i = 0 \end{aligned} \tag{2.9}$$

Thus, the equation for *y* is equivalent to the network model (2.6) and the equation for *y** is equivalent to the spatial-x model (2.7). The errors *u* and ε are assumed to have correlation ρ, which means that the joint distribution of *u* and ε is bivariate normal. According to Breen (1996), this correlation is intrinsic to the model and should be considered inherently immeasurable: "[W]e assume $\rho \neq 0$ in the theoretical model that we posit for the population and not simply for the sample in which we may have omitted the measurement of a variable common to **x** and **s**."

Because the network model (2.6) is selected based on (2.8), estimates of the parameters α and β may be biased and inconsistent. This is because the expected value of y is conditional on $s_i = 1$:

$$E(y_i \mid s = 1, \bar{z}_i, \mathbf{x}_i) = \bar{z}_i \boldsymbol{\alpha} + \mathbf{x}_\mathbf{i} \boldsymbol{\beta} + \rho \sigma_\varepsilon \sigma_u \frac{\phi(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}{\Theta(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}$$

where

$$\frac{\phi(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}{\Theta(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}$$

is the inverse Mill's ratio. In contrast, the expected value of the unconditional, untruncated $y*$ is

$$E(y_i *) = \bar{z}_i \boldsymbol{\alpha} + \mathbf{x}_\mathbf{i} \boldsymbol{\beta}_i$$

The difference between the two expected values means that OLS estimates for the network model (2.6) and spatial-x model (2.7) will generally be different. The two sets of estimates will be the same in two circumstances: One, if ρ is equal to 0, then the two expected values will be the same, which corresponds to the situation in which selection and outcome are independent (Breen, 1996). Two, if the correlation between the estimate of $\frac{\phi(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}{\Theta(\mathbf{x}_{0i}' \boldsymbol{\alpha}_0)}$ and any x variable or $\bar{z}$ is zero, then the OLS estimate of the parameter associated with the x variable or $\bar{z}$ for the network model will be the same as the OLS estimate for the spatial-x model.

Thus, the network model can be considered a special case of the spatial-x model. As a truncated version of the spatial-x model, the network model does not include network observations for the dependent variable and limits network information to independent variables. However, the spatial-x model has two main advantages over the network model. First, the spatial-x model makes use of all the available information (for survey respondents as well as discussants) while preserving the network structure via the spatial-x term. Second, the network model may result in biased and inconsistent estimates due to correlations associated with the selection process. If the correlations are zero, then both the network model and spatial-x model will be appropriate, all else being equal. But if one or more of the correlations are non-zero, then the spatial-x model will be more appropriate. Thus, the spatial-x model is the safer choice.

By understanding the contextual model and network model as special cases of the spatial lag model and spatial-x model, we can see that spatial regression models can do what conventional models of social influence can do and more. While we can estimate the effects of groups on individual-level outcomes with either the contextual model or the spatial lag model, only the latter model explicitly places the individual in a social context (via spatial weights matrix), incorporates the reciprocal causation between individuals and groups, and represents the social influence process as a dynamic process of mutual influence among individuals in a social context. While the effects of social networks on individual-level outcomes can be estimated with either the social network model or the spatial-x model, only the latter makes use of all the available data and does not require assumptions regarding the selection process. Thus, the spatial lag model and the spatial-x model are capable of doing what the contextual model and network model can do, but more.

In particular, the spatial lag model can overcome the problem of reciprocal causation associated with the study of social influence. The spatial lag model explicitly incorporates the reciprocal causation through the spatial lag term; thus, the spatial lag model represents an individual's observed dependent variable as, in part, the outcome of other individuals' observed dependent variables. Because the spatial lag models explicitly incorporate reciprocal causation in this manner, all the covariates become, in a sense, group covariates.

Since the dependent variable y is found on both sides of the equation, any effect from an individual-level covariate will indirectly influence the group as well. Accordingly, the "total effect" of a covariate in spatial lag model consists of direct and indirect effects. For example, suppose an individual is highly interested in politics and that this independent variable is statistically significant. The effect of his political interest will affect his value y for his attitude toward a political issue, such as government spending, but because y is endogenous, his political interest will also (indirectly) affect the observed y for his neighbors as well. In fact, the spatial lag model implies a dynamic data generation process with a feedback loop (Lin et al., 2006). There is no such process represented in the contextual model, social network model, or spatial-x model, which all assume that the dependent variable is the outcome of linear and additive independent variables. This difference is a fundamental advantage of the spatial lag model for the study of social influence. Indeed, the spatial lag model explicitly represents the most plausible process of social influence as mutual influence.

In this chapter, I discussed the theories and models of social influence. I argued that social contexts can be reconceptualized as social spaces that provide opportunities for social interaction and presented three examples: Blau space, ideological space, and geographical space. Following Erbring and Young (1979), I argued that social influence must be understood as mutual influence, which means that each individual influences and is influenced by others in his spatial neighborhood in a social space. I then provided a more formal definition of social influence. Next, I introduced spatial econometrics and showed how spatial regression models can bring together the ideas of social context as social space and social influence as mutual influence. Finally, I examined existing models of social influence and showed that they can be considered as special cases of spatial regression models.

This chapter showed why and how spatial econometrics is well-suited to the study of social influence in politics. It did this by addressing the major shortcomings of conventional models and methods and by providing an alternative framework that overcomes many of those shortcomings.

But what if the researcher wishes simply to capture the role of social influence in a more modest manner? In political science research, researchers may wish to account for social environments, even if they are not interested primarily in modeling social influence directly. In such a circumstance, what does spatial econometrics have to offer over conventional linear regression models? This is the subject of the next chapter, which is an examination of the relationship between linear regression and spatial regression models.

CHAPTER 3: A COMPARISON OF SPATIAL REGRESSION AND LINEAR REGRESSION

Researchers already have a variety of statistical options for dealing with social influence, such as using dummy variables or socio-demographic variables in linear regression models. Why use spatial econometrics instead of these better-known methods? This chapter addresses this question by analyzing the differences and similarities between spatial regression models and linear regression models.

First, I look at the dummy variable regression model and show how dummy variable regression is at best a special case of spatial regression and discuss the substantive advantages of using the spatial regression model for understanding social influence. Next, I turn to a simple regression model with one continuous independent variable and, by analyzing the structure of its hat matrix, show how using this type of model is not a substitute for spatial regression. Third, I extend the discussion of the simple regression model to the general case of multiple regression and show how it compares with spatial regression. Fourth, I show that by discretizing the continuous independent variables, linear regression models can be considered as spatial regression models but only in a very narrow and problematic sense. Finally, I conclude that spatial regression models represent specific processes of mutual influence and contain meaningful parameters for a better understanding of social influence. Consequently, spatial regression models should be preferred over linear regression models for studying social influence.

LINEAR REGRESSION

Since political scientists often study social influence variables using the regression framework, we begin with the classical linear regression model, which is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}$ is an n by 1 vector of dependent variable values, $\mathbf{X}$ is an n by k matrix of independent variables, $\boldsymbol{\beta}$ is a k by 1 vector of coefficients, and $\boldsymbol{\varepsilon}$ is an n by 1 vector of error terms that adhere to the Gaussian-Markov assumptions. The OLS fitted values are given by

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$$

where $\mathbf{b}$ is a k by 1 vector of estimated coefficients given by the well-known formula

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Because of the Gaussian-Markov assumptions regarding the error term, $\mathbf{b}$ is a best linear unbiased estimator of $\boldsymbol{\beta}$.

By combining the equations for the fitted values $\hat{\mathbf{y}}$ and coefficient estimates $\mathbf{b}$, we can see that the fitted values indicated by $\hat{\mathbf{y}}$ are weighted averages of the observed data points indicated by $\mathbf{y}$:

$$\hat{\mathbf{y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

The weights are given by the n by n matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, which is known as the hat matrix. The hat matrix $\mathbf{H}$ is a symmetric and idempotent projection matrix that maps $\mathbf{y}$ into $\hat{\mathbf{y}}$.

We can interpret each element h_{ij} of $\mathbf{H}$ as the amount of influence that each y_j exerts on $\hat{y}_i$, which means we can represent each fitted value $\hat{y}_i$ as follows:

$$\hat{y}_i = \sum_{j=1}^{n} h_{ij} y_j$$

Conventionally, the hat matrix has been used for regression diagnostics. Because its form is comparable to the form of a spatial weights matrix, I will use the hat matrix to show how different linear regressions relate to spatial regression, beginning with the case of dummy variable regression.

DUMMY VARIABLE REGRESSION MODELS VERSUS SPATIAL REGRESSION MODELS

Suppose we have a substantive independent variable such as religion. The possible values for this variable indicate mutually exclusive, exhaustive, and discrete categories; a respondent may be Protestant or Catholic but not both. Membership in each mutually exclusive category can be represented as dummy variables in a regression model. This model is a dummy variable regression model (through the origin) that is equivalent to a one-way analysis of variance (ANOVA) model; it is a type of model that is commonly used for studying group effects. This section will show that the dummy variable regression model is at best a special case of the first-order spatial autoregressive (FAR) model, and because of this, we will see that the FAR model has advantages for the study of social influence.

To represent formally the mutually exclusive groups, consider k groups, where each group comprises n_m members (m = 1, 2, … , k). Let $\mathbf{y}$ be a vector of observations of the dependent variable, $\mathbf{X}$ is a k by k matrix of indicator vectors, $\boldsymbol{\beta}_m$ is a vector of associated coefficients, and $\boldsymbol{\varepsilon}_m$ is a vector of disturbances for the mth group. The dummy variable regression model (through the origin) is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{X}_{k \times k} = \begin{bmatrix} \mathbf{i}_{n_1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{i}_{n_2} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_k} & \mathbf{0}_{n_k} & \cdots & \mathbf{i}_{n_k} \end{bmatrix}$$

and $\mathbf{i}_{n_m}$ is an n_m by 1 vector of ones indicating membership in the mth group and $\mathbf{0}_{n_m}$ is a vector of zeros:

$$\mathbf{i}_{n_m} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \mathbf{0}_{n_m} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Note that each subscript indicates the dimension of each vector. The total sample size is

$$n = n_1 + n_2 + \ldots + n_k$$

Thus, the dummy variable regression model can be written as follows:

$$\begin{bmatrix} \mathbf{y}_{n_1} \\ \mathbf{y}_{n_2} \\ \vdots \\ \mathbf{y}_{n_k} \end{bmatrix} = \begin{bmatrix} \mathbf{i}_{n_1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{i}_{n_2} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{i}_{n_k} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_{n_1} \\ \boldsymbol{\varepsilon}_{n_2} \\ \vdots \\ \boldsymbol{\varepsilon}_{n_k} \end{bmatrix}$$

For an **X** matrix of dummy variables, the hat matrix **H** turns out to be a block-diagonal matrix, which is a square diagonal matrix whose diagonal elements are square matrices and the off-diagonal elements are zero matrices. To see this, we take advantage of the fact that

$$\mathbf{X}'\mathbf{X} = diag(n_1, n_2, \ldots, n_k)$$

so that

$$
\begin{aligned}
\mathbf{H} &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\
&= \begin{bmatrix} \mathbf{i}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{i}_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{i}_{n_k} \end{bmatrix} \cdot diag\left(\frac{1}{n_1}, \frac{1}{n_2}, \ldots, \frac{1}{n_k}\right) \cdot \begin{bmatrix} \mathbf{i}'_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{i}'_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{i}'_{n_k} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\mathbf{i}_{n_1}\mathbf{i}_{n_1}'}{n_1} & 0 & \cdots & 0 \\ 0 & \frac{\mathbf{i}_{n_2}\mathbf{i}_{n_2}'}{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\mathbf{i}_{n_k}\mathbf{i}_{n_k}'}{n_k} \end{bmatrix} \\
&= diag\left(\frac{\mathbf{i}_{n_1}\mathbf{i}_{n_1}'}{n_1}, \frac{\mathbf{i}_{n_2}\mathbf{i}_{n_2}'}{n_2}, \ldots, \frac{\mathbf{i}_{n_m}\mathbf{i}_{n_m}'}{n_m}\right) \\
&= \oplus_{m=1}^{k}\left(\frac{\mathbf{i}_{n_m}\mathbf{i}_{n_m}'}{n_m}\right)
\end{aligned}
$$

where $\oplus$ is the conventional notation for the matrix direct sum. We can see that the n by n matrix $\mathbf{H}$ is a block-diagonal matrix where the mth diagonal block is the n_m by n_m matrix $\frac{1}{n_m}\mathbf{i}_{n_m}\mathbf{i}_{n_m}'$. This is a matrix whose elements are all equal to $1/\,n_m$. It is easy to see that the fitted values are simply the arithmetic means of each group:

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$$

$$= \begin{bmatrix} \frac{\mathbf{i}_{n_1}\mathbf{i}_{n_1}'}{n_1} & 0 & \cdots & 0 \\ 0 & \frac{\mathbf{i}_{n_2}\mathbf{i}_{n_2}'}{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\mathbf{i}_{n_k}\mathbf{i}_{n_k}'}{n_k} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{y}_{n_1} \\ \mathbf{y}_{n_2} \\ \vdots \\ \mathbf{y}_{n_k} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{\mathbf{y}}_{n_1} \\ \bar{\mathbf{y}}_{n_2} \\ \vdots \\ \bar{\mathbf{y}}_{n_k} \end{bmatrix}$$

Note that $\bar{\mathbf{y}}_{n_m}$ is a vector in which the elements are the mean of the elements of $\mathbf{y}_{n_m}$, *i.e.,* the element $\bar{y}_{n_m}$ is the mean value of y for the n_m members of the mth group. This means that $\hat{\mathbf{y}}_i$ is the mean of the y values of group m associated with individual i.

Now that we have uncovered the structure of hat matrix $\mathbf{H}$ for dummy variable regression models, we will next see that $\mathbf{H}$ bears a strong resemblance to a specific type of (standardized) spatial weights matrix $\mathbf{W}$. In fact, the structures of $\mathbf{H}$ and $\mathbf{W}$ turn out to be very similar for *disjoint neighborhoods* (*i.e.*, mutually exclusive groups); using this similarity, the following discussion will show that when n_m becomes very large,

$$\mathbf{Xb} = \mathbf{Hy} \approx \mathbf{Wy}$$

We will then use this asymptotic equivalence to show that the dummy variable regression model is a special case of the FAR model.

Under the spatial regression framework, group membership is represented in a spatial weights matrix $\mathbf{W}$. Spatial weights matrices are based on the researcher's definition of what a *neighborhood* (*i.e.*, group) is; for example, neighborhoods can be based on religion, in the sense that two individuals are "neighbors" if they belong to the same religion. When neighborhoods are "disjoint," two individuals are neighbors if they reside in the same neighborhood but not if they reside in different neighborhoods. Disjoint neighborhoods in the spatial regression framework are analogous to the mutually exclusive categories used in the dummy variable regression model (through the origin).

A spatial weights matrix representing disjoint neighborhoods is an n by n matrix $\mathbf{W}^*$, where each element $w^*_{ij} = 1$ when individuals i and j are neighbors and $w^*_{ij} = 0$ when $i = j$ or when individuals i and j are not neighbors. Furthermore, each diagonal element w^*_{ii} is equal to 0, under the reasoning that individuals are not their own neighbors. Conventionally, the spatial weights matrix is row-standardized so that the elements of each row sum to 1. In this chapter, $\mathbf{W}$ will denote the row-standardized version of $\mathbf{W}^*$. Note that post-multiplying $\mathbf{W}$ by the dependent variable vector $\mathbf{y}$ yields the spatial lag $\mathbf{Wy}$, where each element of $\mathbf{Wy}$ represents the average y value of observation i's neighbors.

We have seen that in a dummy variable regression model (through the origin), the hat matrix $\mathbf{H}$ is a block diagonal matrix where the mth diagonal block is the n_m by n_m matrix

$$\frac{1}{n_m}\mathbf{i}_{n_m}\mathbf{i}_{n_m}'$$

Similarly, in spatial regression model with disjoint neighborhoods, a spatial weights matrix $\mathbf{W}$ representing the k disjoint neighborhoods is also a block diagonal matrix. As Lin et al (2006) observed,

> A characteristic of such disjoint neighborhoods is that the spatial weights matrix, $\mathbf{W}$, can always be arranged in a "block-diagonal" fashion. That is, there will be blocks of square matrices with non-zero off-diagonal elements along the principal diagonal line of $\mathbf{W}$, with each block representing a neighborhood while all other elements are zero outside these blocks.

In the following discussion, we will establish that the diagonal blocks of the block diagonal matrix $\mathbf{W}$ are asymptotically equivalent to the diagonal blocks of $\mathbf{H}$.

Suppose there are k neighborhoods and each neighborhood contains n_m individuals (m = 1, 2, … , k). This is the same situation as the one discussed in the case of dummy variable regression, but here we use the spatial econometric term "neighborhood" instead of group.

Let **W*** be an unstandardized spatial weights matrix such that $w^*_{ij} = 1$ when individuals i and j are neighbors and $w^*_{ij} = 0$ when $i = j$ or when individuals i and j are not neighbors. **W*** is the following block-diagonal matrix:

$$\mathbf{W}^*_{n\times n} = \begin{bmatrix} \mathbf{i}_{n_1}\mathbf{i}_{n_1}'-\mathbf{I}_{n_1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{i}_{n_2}\mathbf{i}_{n_2}'-\mathbf{I}_{n_2} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_k} & \mathbf{0} & \cdots & \mathbf{i}_{n_k}\mathbf{i}_{n_k}'-\mathbf{I}_{n_k} \end{bmatrix}$$

where each block $\mathbf{i}_{n_m}\mathbf{i}_{n_m}'-\mathbf{I}_{n_m}$ is a n_m by n_m matrix representing the mth neighborhood, minus individual i (since individual i is not his own neighbor). To row-standardize **W***, we divide each row by the number of neighbors (n_m - 1), which is a scalar. Letting **W** represent the resulting row-standardized spatial weights matrix:

$$\mathbf{W}_{n\times n} = \begin{bmatrix} \dfrac{\mathbf{i}_{n_1}\mathbf{i}_{n_1}'-\mathbf{I}_{n_1}}{n_1-1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \dfrac{\mathbf{i}_{n_2}\mathbf{i}_{n_2}'-\mathbf{I}_{n_2}}{n_2-1} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_k} & \mathbf{0} & \cdots & \dfrac{\mathbf{i}_{n_k}\mathbf{i}_{n_k}'-\mathbf{I}_{n_k}}{n_k-1} \end{bmatrix}$$

The spatial lag of **y** is **Wy**, which is given by

$$\mathbf{Wy} = \begin{bmatrix} \frac{\mathbf{i}_{n_1}\mathbf{i}_{n_1}'-\mathbf{I}_{n_1}}{n_1-1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \frac{\mathbf{i}_{n_2}\mathbf{i}_{n_2}'-\mathbf{I}_{n_2}}{n_2-1} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_k} & \mathbf{0} & \cdots & \frac{\mathbf{i}_{n_k}\mathbf{i}_{n_k}'-\mathbf{I}_{n_k}}{n_k-1} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{y}_{n_1} \\ \mathbf{y}_{n_2} \\ \vdots \\ \mathbf{y}_{n_k} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{n_1}{n_1-1}\left(\bar{\mathbf{y}}_{n_1} - \frac{\mathbf{y}_{n_1}}{n_1}\right) \\ \frac{n_2}{n_2-1}\left(\bar{\mathbf{y}}_{n_2} - \frac{\mathbf{y}_{n_2}}{n_2}\right) \\ \vdots \\ \frac{n_k}{n_k-1}\left(\bar{\mathbf{y}}_{n_k} - \frac{\mathbf{y}_{n_k}}{n_k}\right) \end{bmatrix}$$

But note that when n_m approaches infinity,

$$\frac{n_m}{n_m-1} \to 1; \quad \frac{1}{n_m} \to 0$$

so that

$$\mathbf{Wy} = \begin{bmatrix} \frac{n_1}{n_1-1}\left(\bar{\mathbf{y}}_{n_1} - \frac{\mathbf{y}_{n_1}}{n_1}\right) \\ \frac{n_2}{n_2-1}\left(\bar{\mathbf{y}}_{n_2} - \frac{\mathbf{y}_{n_2}}{n_2}\right) \\ \vdots \\ \frac{n_k}{n_k-1}\left(\bar{\mathbf{y}}_{n_k} - \frac{\mathbf{y}_{n_k}}{n_k}\right) \end{bmatrix} \approx \begin{bmatrix} \bar{\mathbf{y}}_{n_1} \\ \bar{\mathbf{y}}_{n_2} \\ \vdots \\ \bar{\mathbf{y}}_{n_k} \end{bmatrix}$$

Thus,

$$\mathbf{Wy} \approx \mathbf{Hy} = \begin{bmatrix} \bar{\mathbf{y}}_{n_1} \\ \bar{\mathbf{y}}_{n_2} \\ \vdots \\ \bar{\mathbf{y}}_{n_k} \end{bmatrix}$$

The asymptotical equivalence of $\mathbf{Wy}$ and $\mathbf{Hy}$ means that the spatial lag of *y* and the OLS fitted values of a dummy variable regression model are also asymptotically equivalent:

$$\hat{\mathbf{y}} = \mathbf{Hy} \approx \mathbf{Wy}$$

This is a special case of the FAR model $\hat{\mathbf{y}} = \hat{\rho}\mathbf{Wy}$ when the spatial autoregressive parameter $\hat{\rho} = 1$. Thus, the dummy variable regression model (through the origin) is a special case of a spatial regression model, when the neighborhoods (or groups) are exclusive, exhaustive, and discrete.[10]

However, the difference between the FAR model and the dummy variable regression model is not a trivial one. First, by restricting the spatial autoregressive parameter ρ to 1, the dummy variable regression model forces the assumption that individual observations are essentially the same as their group averages. Under the dummy variable framework, the effect of the group average is always equal to 1; under the FAR framework, the effect of the group average can take on other values. For the purposes of studying social influence, the latter is clearly more relevant, since estimating the effect of the group average *is* estimating the effect of the group on the individual.

Second, the dummy variables account only for influence in one direction, while the FAR model can measure mutual influence. The presence of the *y* variable on both sides of the FAR model represents an endogenous process in which observations influence and are influenced by neighboring observations. In the study of social influence, the idea that individuals *influence* and *are influenced by* other individuals in their groups implies endogeneity. While the FAR model directly models this endogeneity among neighboring observations, it is absent in the dummy variable regression model. In the dummy variable regression model, the dependent variable (such as a political attitude) is simply a linear combination of group-

[10] Lin et al. (2006) reached the same conclusion using a slightly different method.

membership indicators. It implies that groups affect individuals but, paradoxically, individuals do not affect the groups to which they belong.

SIMPLE LINEAR REGRESSION MODELS VERSUS SPATIAL REGRESSION MODELS

In addition to dummy variable regression models, researchers can assess social influence by using a linear regression model with continuous independent variables that indicate social characteristics. While this type of model is not usually found in studies that are explicitly about social influence, researchers in other areas of study often find it useful to include in their models independent variables such as income or education – variables that can be interpreted as individual *or* social characteristics. For example, household income may be considered an individual-level indicator of material resources or an indicator of socio-economic status. When independent variables are interpreted as social characteristics, the models that employ them necessarily make an implicit claim about social influence (*e.g.*, group interests). These implicit claims of social influence fall short of what can be accomplished in spatial regression models.

In this section, I show that a simple linear regression model with one continuous variable is *not* a special case of a spatial regression model. In contrast to the dummy variable regression model (through the origin), the simple linear regression model does not have a hat matrix **H** that is comparable to a spatial weights matrix **W**; because of this, the estimates from a linear regression model are not equivalent to the estimates from a corresponding spatial regression model, which are more relevant for studying social influence.

The classical linear regression model is once again given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}$ is an n by 1 vector of dependent variable values, $\boldsymbol{\beta}$ is a k+1 by 1 vector of coefficients, and $\boldsymbol{\varepsilon}$ is an n by 1 vector of error terms that adhere to the Gauss-Markov assumptions. Let $\mathbf{X}$ be the following n by 2 matrix representing one continuous independent variable and an intercept:

$$\mathbf{X}_{n\times 2} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

The OLS fitted values are given by

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$$

where $\mathbf{H}$ is equal to $\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$. It has already been shown that for disjoint neighborhoods, the hat matrix $\mathbf{H}$ and spatial weights matrix $\mathbf{W}$ are asymptotically equivalent. But in our simple linear regression model, $\mathbf{H}$ is based on a continuous variable (rather than a set of dummy variables). We will see that $\mathbf{H}$ is not equivalent to a spatial weights matrix $\mathbf{W}$ that is based on that continuous variable.

Under the spatial regression framework, a spatial weights matrix $\mathbf{W}$ that is based on a continuous variable can be considered an inverse distance matrix. In an inverse distance matrix $\mathbf{W*}$, the elements w_{ij} are functions of the inverse distance between two observations i and j. Each element w_{ij} is larger when observations i and j are closer together and w_{ij} is smaller when observations i and j are far apart. This structure of $\mathbf{W}$* has substantive consequences for the meaning of the spatial lag $\mathbf{W}$*$\mathbf{y}$. The ith element of $\mathbf{W}$*$\mathbf{y}$ is given by

$$\sum_{j\neq i} w_{ij} y_j$$

when each y_j is given more weight when w_{ij} is large and less weight when w_{ij} is small. Since w_{ij} is large when the distance between observations i and j is small, and w_{ij} is small when the distance between observations *i* and *j* is large, this means that the spatial lag **W*y** is the weighted sum of observation *i*'s neighbors' *y* values, where neighbors that are farther away count for less and neighbors that are closer by count for more.

To assess whether **H** is a type of inverse distance matrix, we need to determine whether each h_{ij} is like w_{ij} (for $i \neq j$) in the sense that they are both based on the distance between observations *i* and *j*. If so, then we can classify **H** as a type of inverse distance matrix. But as we will see, we cannot.

Determining whether each h_{ij} is based on the distance between observations *i* and *j* is not a straightforward task because each h_{ij} (for $i \neq j$) is defined relative to a reference point: it is the leverage of y_j on $\hat{y}_i$. That means we need to employ additional points for comparison. Accordingly, let $d_{ij} = \|x_j - x_i\|$ represent the "distance" between observations *j* and *i*, and let $d_{ik} = \|x_k - x_i\|$ represent the distance between observations *k* and *i*. We will compare observations *j* and *k* and see if their leverages on $\hat{y}_i$ correspond with d_{ij} and d_{ik}. If **H** is a type of distance-based matrix, then a difference between d_{ij} and d_{ik} should correspond to a difference between h_{ij} and h_{ik}. Specifically, if h_{ij} is a spatial weight, it should be small when the distance between observations *j* and *i*, is large, and it should be large when the distance between observations *j* and *i* is small. Accordingly, if $d_{ij} > d_{ik}$, we should expect $h_{ij} < h_{ik}$, and if $d_{ij} < d_{ik}$, we should expect $h_{ij} > h_{ik}$.

The **X** matrix for a simple linear regression model implies that the *n* by *n* hat matrix **H** has the following general structure:

$$\mathbf{H} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$$

$$= \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \cdot \begin{bmatrix} \dfrac{\sum_i x_i^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \dfrac{-\sum_i x_i}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \\ \dfrac{-\sum_i x_i}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \dfrac{n}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\sum_i x_i^2 - 2nx_1\overline{x} + nx_1^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \dfrac{\sum_i x_i^2 - n(x_1 + x_2)\overline{x} + nx_1x_2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \cdots & \dfrac{\sum_i x_i^2 - n(x_1 + x_n)\overline{x} + nx_1x_n}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \\ \dfrac{\sum_i x_i^2 - n(x_1 + x_2)\overline{x} + nx_1x_2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \dfrac{\sum_i x_i^2 - 2nx_2\overline{x} + nx_2^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \cdots & \dfrac{\sum_i x_i^2 - n(x_2 + x_n)\overline{x} + nx_1x_n}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\sum_i x_i^2 - n(x_1 + x_n)\overline{x} + nx_1x_n}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \dfrac{\sum_i x_i^2 - n(x_2 + x_n)\overline{x} + nx_2x_n}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} & \cdots & \dfrac{\sum_i x_i^2 - 2nx_n\overline{x} + nx_n^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \end{bmatrix}$$

From the matrix above, we can see that each off-diagonal element h_{ij} is given by the following:

$$h_{ij} = \frac{\sum_i x_i^2 - (x_i + x_j)n\overline{x} + nx_i x_j}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

$$= \frac{n\left[\frac{1}{n}\sum_i (x_i - \overline{x})^2 + (x_i - \overline{x})(x_j - \overline{x})\right]}{n\sum_i (x_i - \overline{x})^2}$$

$$= \frac{1}{n} + \frac{(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2}$$

Note that when either x_i or x_j are equal to the global mean $\overline{x}$, h_{ij} is equal to n^{-1}. The off-diagonal element h_{ij} is commonly interpreted as the amount of influence exerted on $\hat{y}_i$ by y_j.

Each diagonal element of **H** is given by the following:

$$h_{ii} = \frac{\sum_i x_i^2 - 2x_i n\overline{x} + nx_i^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2}$$

$$= \frac{\frac{1}{n}\left[\sum_i (x_i - \overline{x})^2 + (x_i - \overline{x})^2\right]}{\sum_i (x_i - \overline{x})^2}$$

$$= \frac{1}{n} + \frac{(x_i - \overline{x})^2}{n\sum_i (x_i - \overline{x})^2}$$

It is known that $0 \le h_{ii} \le 1$, since **H** is a symmetric idempotent matrix (see Hoaglin and Welsch, 1978). Note once again that when x_i is equal to the global mean $\overline{x}$, h_{ii} is equal to n^{-1}. The diagonal element h_{ii} is the amount of influence exerted on $\hat{y}_i$ by y_i.

Let Δ denote the difference between h_{ij} and h_{ik}. Then

$$\Delta = h_{ij} - h_{ik}$$
$$= \left[\frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \right] - \left[\frac{1}{n} + \frac{(x_i - \bar{x})(x_k - \bar{x})}{\sum_i (x_i - \bar{x})^2} \right]$$
$$= \frac{(x_i - \bar{x})(x_j - x_k)}{\sum_i (x_i - \bar{x})^2}$$

Whether Δ is positive or negative depends on $x_j - x_k$ as well as $x_i - \bar{x}$. Using this fact, we can see that $d_{ij} > d_{ik}$ does not consistently correspond with $h_{ij} < h_{ik}$, and $d_{ij} < d_{ik}$ does not correspond with $h_{ij} > h_{ik}$.

For example, suppose $x_i > \bar{x} > x_j > x_k$. This means that $h_{ij} > h_{ik}$ and that $d_{ij} < d_{ik}$. Accordingly, the conditions for **H** to be a spatial weights matrix are satisfied.

But suppose $\bar{x} < x_i < x_j < x_k$. This would mean that $h_{ij} < h_{ik}$ and that $d_{ij} < d_{ik}$. In order for **H** to be a spatial weights matrix, however, we require that $d_{ij} > d_{ik}$. Thus, the conditions for **H** to be a spatial weights matrix are no longer satisfied. This counterexample is sufficient grounds to conclude that in the case of simple linear regression, the hat matrix is not a type of spatial weights matrix.

MULTIPLE LINEAR REGRESSION MODELS VERSUS SPATIAL REGRESSION MODELS

It has been shown that for disjoint neighborhoods, the dummy variable regression model (through the origin) is a special case of the FAR model (with $\rho = 1$). It has also been shown that this equivalence does not extend to a simple regression model with one continuous independent variable and first-order spatial autoregressive model with a distance spatial weights matrix based on the same continuous variable. This section generalizes the results for the simple regression model; we will see that the equivalence also does not hold for a multivariate regression model with *k* continuous variables.

The classical linear regression model is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}$ is an n by 1 vector of dependent variable values, $\boldsymbol{\beta}$ is a k +1 by 1 vector of coefficients, and $\boldsymbol{\varepsilon}$ is an n by 1 vector of error terms that adhere to the Gaussian-Markov assumptions. Let $\mathbf{X}$ be the following n by k+1 matrix representing k continuous independent variable and an intercept:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{kn} \end{bmatrix}$$

The OLS fitted values are given by

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$$

where $\mathbf{H}$ is equal to

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

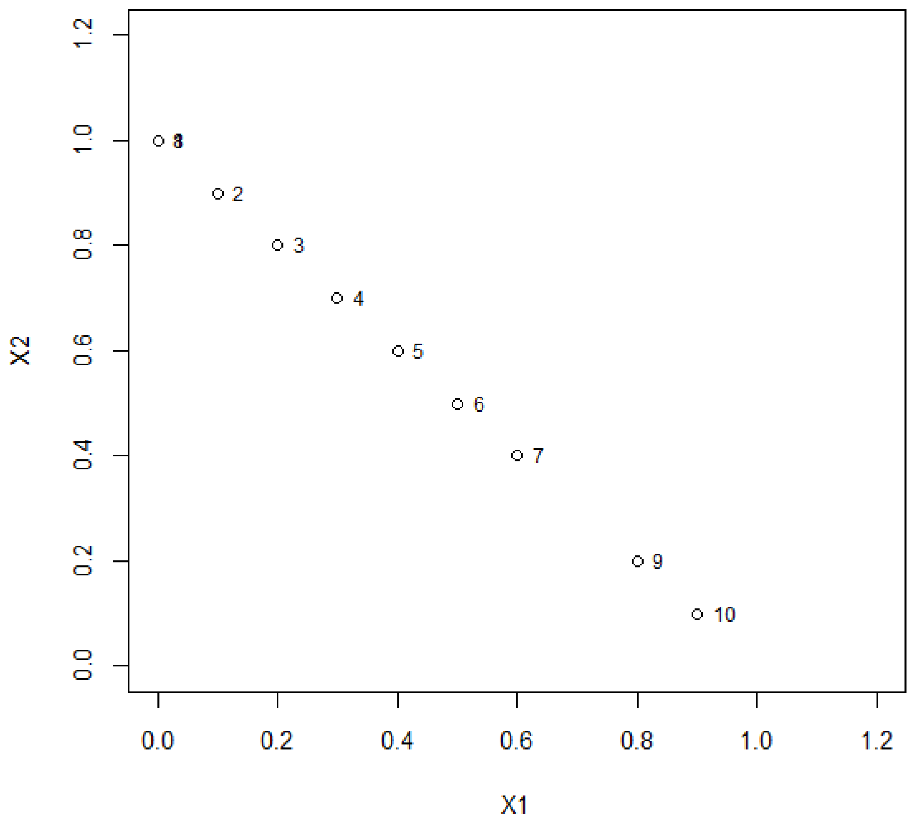
To demonstrate why a multivariate regression with k continuous variables is not equivalent to a spatial regression model with corresponding spatial weights matrix, we turn to a numerical example, since the form of the hat matrix $\mathbf{H}$ is too complicated. We have the following simulated observations of two continuous independent variables for ten respondents:

Table 3.1. 10 observations

RESPONDENTS	X1	X2
1	0.0	1.0
2	0.1	0.9
3	0.2	0.8
4	0.3	0.7
5	0.4	0.6
6	0.5	0.5
7	0.6	0.4
8	0.0	1.0
9	0.8	0.2
10	0.9	0.1

To examine the distance (or similarity) between the respondents, we can plot these data points on a map (Figure 3.1). Note that each point is associated with a respondent number.

Figure 3.1. 10 Observations of Simulated Data



The corresponding hat matrix **H** is given by

Table 3.2. Hat Matrix Based on Two Continuous Independent Variables

	1	2	3	4	5	6	7	8	9	10
1	0.26	0.22	0.17	0.13	0.09	0.05	0.01	0.26	-0.07	-0.12
2	0.22	0.19	0.16	0.12	0.09	0.06	0.03	0.22	-0.03	-0.06
3	0.17	0.16	0.14	0.12	0.10	0.08	0.06	0.17	0.02	0.00
4	0.13	0.12	0.12	0.11	0.10	0.09	0.08	0.13	0.06	0.05
5	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.09	0.11	0.11
6	0.05	0.06	0.08	0.09	0.10	0.12	0.13	0.05	0.16	0.17
7	0.01	0.03	0.06	0.08	0.10	0.13	0.15	0.01	0.20	0.22
8	0.26	0.22	0.17	0.13	0.09	0.05	0.01	0.26	-0.07	-0.12
9	-0.07	-0.03	0.02	0.06	0.11	0.16	0.20	-0.07	0.29	0.34
10	-0.12	-0.06	0.00	0.05	0.11	0.17	0.22	-0.12	0.34	0.40

Using the same reasoning we used to examine the simple linear regression model, we can use the elements of **H** in Table 3.2 to assess whether the difference between d_{ij} and d_{ik} corresponds to the difference between h_{ij} and h_{ik}. If **H** is indeed a kind of distance-based matrix, then a greater difference between d_{ij} and d_{ik} should correspond to the smaller difference between h_{ij} and h_{ik}.

Consider respondent 4 as a baseline and respondents 5 and 6 for comparison. From Figure 3.1, we can see that $d_{45} < d_{46}$, and from Table 3.2, we can see that $h_{45} > h_{46}$. Since a greater distance corresponds with a lesser hat value, this example is consistent with the idea that **H** is a spatial weights matrix.

Now let us look at respondent 4 as a baseline and respondents 2 and 5 for comparison. From Figure 2.1, we can see that $d_{42} > d_{45}$. But from Table 3.2, we can see that $h_{42} > h_{45}$, which means that we have an example of a greater distance corresponding with a *greater* hat value rather than a *smaller* hat value.

Based on this numerical example, we can see that a greater (lesser) difference between d_{ij} and d_{ik} does not reliably correspond with a lesser (greater) difference between h_{ij} and h_{ik}. Thus, we can conclude that **H** is not a type of distance-based spatial weights matrix.

As in the case of the simple linear regression model, the multivariate linear regression model is not equivalent to the spatial regression model. A multivariate regression model does not reflect the relationship between influence and distance the way that a spatial regression model can; and it is the ability to capture this relationship that makes the latter model better for studying social influence.

DISCRETIZED LINEAR REGRESSION MODELS VERSUS SPATIAL REGRESSION MODELS

From the previous sections, we see that the hat matrices for the simple regression model and multivariate regression model do not behave like distance-based spatial weights matrices in spatial regression models. Because of this, we can conclude that linear regression models are generally not special cases of spatial regression models. However, this conclusion is only a general one because linear regression models *can* actually be expressed as spatial regression models. By discretizing the continuous independent variables, we can transform any linear regression model into a dummy variable regression model (through the origin), which is a special case of the FAR model.

The classical linear regression model is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}$ is an n by 1 vector of dependent variable values, $\boldsymbol{\beta}$ is a vector of coefficients, and $\boldsymbol{\varepsilon}$ is an n by 1 vector of error terms that adhere to the Gaussian-Markov assumptions. Let $\mathbf{X}$ be the following n by k matrix representing k continuous independent variables:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}$$

Each of the *k* continuous variables can be discretized into *p* intervals, where *p* is an integer and $p \geq 2$.[11] In other words, we represent the *m*th continuous variable *x* as a set of *p* dummy variables:

$$x_m \equiv \{d_{m1}, d_{m2}, \ldots, d_{mp}\}$$

For example, if x_1 were age in years, d_{11} might be children (ages 12 and under), d_{12} might be teenagers (ages 13 through 17), d_{13} might be young adults ages (18 through 34), d_{14} might be mature adults ages (35 through 65), and d_{15} might be the elderly (ages 65 and over). For a respondent who is 45 years old, we would say that $d_{14} = 1$ and the remaining dummy variables would all be set to 0.

For a simple linear regression model with one continuous variable (*k* = 1), discretizing the continuous independent variable simply means transforming the simple linear regression model into a dummy variable regression (through the origin) with *p* exhaustive and mutually exclusive dummy variables. Since no respondent can score a "1" for more than one dummy variable, we can think of each dummy variable as a disjoint neighborhood. As we saw in a previous section, disjoint neighborhoods can be represented in a spatial weights matrix (with zeroes on the diagonal), which means that a dummy variable regression (through the origin) is a special case of a FAR model (when $\rho = 1$). In other words, by discretizing the continuous variable, a simple linear regression model can be considered a special case of a spatial regression model.

[11] For the sake of simplicity, we assume that *p* is the same for all *k* variables.

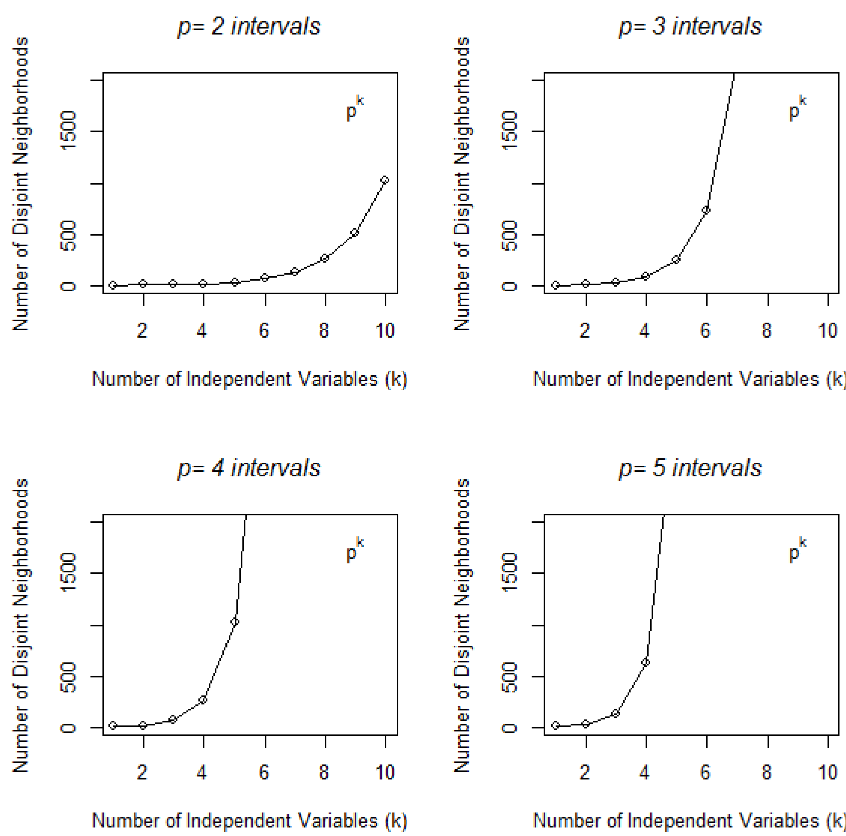
For a multivariate linear regression with $k > 1$ continuous variables, we can discretize all the continuous independent variables into p intervals and define each neighborhood as a unique combination of the kp discretized variables. Accordingly, the number of dummy variables for the discretized multivariate linear regression model is p^k, since there are p possible intervals for each of the k variables. By allowing each respondent to belong to only one of the p^k neighborhoods, the neighborhoods are necessarily disjoint and can be expressed as a spatial weights matrix. The asymptotical equivalence between a dummy variable regression (through the origin) based on the p^k neighborhoods and the FAR model (when $\rho = 1$) follows from this. In this way, a multivariate linear regression model can be considered a special case of a spatial regression model.

While any linear regression model can be expressed as a spatial regression model, it would generally be unwise to do so because of how quickly the number of dummy variables can accumulate during the discretization process. To see how this happens, recall that each continuous independent variable can be discretized into $p \geq 2$ intervals. In the simplest case, each continuous independent variable is dichotomized, which means that $p = 2$. As k (the number of independent variables) increases, the number of dummy variables required for the discretized linear regression increases exponentially. For example, by increasing the number of independent variables from 4 to 5, the number of dummy variables jumps from 16 to 32 and by increasing the number of independent variables from 5 to 6, the number of dummy variables jumps from 32 to 64.

The increase in the number of dummy variables (and subsequent loss of degrees of freedom) is more rapid when we also allow p to increase (see Figure 3.2). For example, consider $k = 3$ independent variables and let us discretize those variables into $p = 4$

intervals. Despite the modesty of those numbers, we end up with a daunting 64 dummy variables.

Figure 3.2. Number of Disjoint Neighborhoods (Dummy Variables) for Different Intervals (p) and Continuous Variables (k)



While the discretized linear regression model is a special case of a spatial regression model, it is a problematic one. The fact that any linear regression model *can* be converted into a spatial regression model via discretization does not mean that it *should* be done in actual data analyses. First, the rapid increase in the number of dummy variables at the higher values of p and k makes such a conversion impractical. Second, the discretization process results in a significant loss of information. The discretized linear regression model implies completely separate means for each group, even if some groups are more similar than others. This is because the discretized linear regression model is essentially an ANOVA model, and ANOVA models make no assumptions regarding the statistical relationship between the independent and dependent variables. Consequently, in a linear regression model, respondents with similar independent variable values will have similar estimated dependent variable values and there will be a linear relationship between the independent and dependent variable. In contrast, in a discretized linear regression model or ANOVA model, respondents with similar independent variables values are forced to have completely different means; there is no recognition of similarity among the groups and there is no relationship among the means. Thus, the discretized linear regression model is a rather crude version of the linear regression model.

The usefulness of the discretized linear regression model lies not in its application, but in its connection to the spatial regression model. It is a theoretical device that allows us to compare and contrast the application of the general linear regression model and the spatial regression model to social influence analyses.

When it comes to linear regression models, a regression coefficient is the effect of a unit increase in one independent variable, all else being equal. In political science research, however, we are often not interested in isolating the effect of a single independent variable but in understanding the effect of a combination of related variables. For example, consider the attitude toward government spending as the dependent and consider age, income, and education as the (roughly) continuous independent variables. While the effects of unit increases in age, income, or education may be interesting and statistically significant in their own right, it is often more meaningful to consider combinations of these characteristics: What is the effect of being poor, young, and college-educated on one's attitude toward government spending? What is the effect of being affluent, middle-aged, and high-school-educated? Estimating these kinds of combined effects is important.

To estimate combinations of effects on the dependent variable, researchers can simply add up different combinations of estimated effects from a linear regression model. By conceptualizing a discretized linear regression model (i.e., a dummy variable regression (through the origin) with p^k dummy variables), we can see that the combined effects are equivalent to the group means, such as the average political attitude of all poor, young, and college-educated individuals and the average political attitude of all affluent, middle-aged high-school-educated individuals.

But as we saw in an earlier section comparing general dummy variable regression and spatial regression, group means can be considered as social influence only in a very narrow sense. If we think of social influence as the effect of belonging to a group, using a group mean requires the assumption that individuals who belong to a group fully assimilate the group mean. As a special case of the FAR model, the discretized linear regression model illustrates this assumption by restricting the spatial autoregressive parameter ρ to 1. The constraint $\rho = 1$ means that the social influence completely determines an individual's observed y value such that

$$\hat{y}_i = \bar{y}_m$$

This means that the group effect is *equal* to the group mean. By using the FAR model, we acknowledge ρ as a parameter that can be estimated, allowing the group effect to be *proportional* to the group mean:

$$\hat{y}_i = \rho \cdot \bar{y}_m$$

Because the spatial autoregressive parameter ρ is the *effect* of the group mean, it can more accurately be called a measure of social influence.

CONCLUSION

Recognizing the relevance of groups and social characteristics, political scientists usually include them as control variables in their regression models. Political methodologists have often criticized this "garbage can" approach for lacking any real explanatory power. In this spirit, Achen (2005) writes:

> Countries, wars, racial categories, religious preferences, education levels, and other variables that change people's coefficients are "controlled" with dummy variables that are completely inadequate to modeling their effects. The result is a long list of independent variables, a jumbled bag of nearly unrelated observations, and often a hopelessly bad specification with meaningless (but statistically significant with several asterisks!) results.

If demographic, ideological, and geographical variables are important for understanding political attitudes and if Achen and his partisans are correct, then having them as control variables in a linear regression model would not help us say very much about the role of social influence in political attitude formation. Doing so would not capture the most theoretically interesting aspect of these characteristics – that race/ethnicity, age, religion, education, occupation, and gender provide contexts in which individuals think and talk about politics.

In this chapter, I showed that spatial regression models offer meaningful advantages over linear regression models. First, by demonstrating how the dummy variable regression model is a special case of the first-order spatial autoregressive (FAR) model, I showed that the latter model offer advantages over the former. The FAR model contains a substantively useful parameter (ρ) that quantifies the amount of social influence exerted by social groups. The FAR model also directly incorporates a dynamic, endogenous process of influence that dummy variable regression and linear regression models lack.

Next, I showed that simple and multiple linear regression models are generally not special cases of spatial regression models by demonstrating that their hat matrices do not behave like the distance-based spatial weights matrices used in spatial regression models. These results explain why including a collection of geographic, demographic, and ideological control variables in a linear regression will not produce the same results as including those same variables in a spatial weights matrix in a spatial regression model.

Finally, I discussed how simple and multiple linear regression models can be forced into special cases of spatial regression models by discretizing the continuous independent variables. Since this procedure produced a dummy variable regression model that quickly lost degrees of freedom even at modest numbers of independent variables, the discretized linear regression was theoretically useful for showing how linear regression models and spatial regression models can be used to address the same questions of group influence. But as in the case of the dummy variable regression model, the discretized linear regression model falls short of the spatial regression model's ability to model individuals with respect to their groups.

Because spatial regressions are much more meaningful than dummy variable regression models and linear regression models, spatial econometrics can serve as an alternative to the much-maligned garbage-can approach.

CHAPTER 4: EMPIRICAL RESEARCH DESIGN

Up to this point, this dissertation has been mainly theoretical. I have looked at the methodological problem of reciprocal causation, proposed an alternative framework based on spatial econometric concepts, reframed the idea of social context as social space, emphasized the idea of social influence as mutual influence, introduced spatial econometric models of social influence, and compared and contrasted spatial regression models with more conventional regression models. The utility of these theoretical considerations remains to be illustrated in actual empirical applications of such a spatial econometric approach to understanding social influence, which are the subjects of chapters 5, 6 and 7. This chapter describes the research design of these applications.

RESEARCH QUESTION

This dissertation focuses on how researchers can better analyze and understand social influence in politics by using a spatial econometric framework. Motivated by the problems of reciprocal causation and social distance, this dissertation is concerned with the following major question and related corollary questions:

1. *Main Question*: What is the best way to study social influence on political attitudes?
2. *Secondary Questions*:

a. How can social contexts be reconceptualized as spaces for spatial econometric analysis? What are the dimensions of these spaces, and how should they be used in spatial regression models?

b. Compared to the dominant models of social influence, what advantages and disadvantages does a spatial econometric approach offer?

c. To what extent are demographic, ideological, and geographical contexts important for understanding social influence?

UNIT OF ANALYSIS

The unit of analysis is the individual, and the phenomena under study are individual-level political attitudes toward various political issues.

DATA SOURCE

This dissertation relies on the 2004 American National Election Studies dataset for analysis. This is an unusual (and unprecedented) choice because studies of social influence typically rely on social network data or a combination of datasets (such as individual-level survey data and census data) in order to obtain the individual-level and group-level data necessary for a contextual analysis or a social network analysis. In contrast, the spatial econometric analyses in this dissertation are based on individual-level data from a well-known mass survey. But how can a study of social influence be based on a public opinion survey that consists of randomly selected individuals who have probably never met, never interacted, and therefore never *actually* influenced each other?

The justification for using a public opinion survey lies in *structural equivalence*, a concept found in network analysis. Structural equivalence is the idea that actors who have identical (or sufficiently similar) relations to all other actors in a social system can be considered equivalent (see Lorrain and White, 1971; Sailer, 1978; Burt, 1987). Burt (1987) provides the following definition:

> Structurally equivalent people occupy the same position in social structure and so are proximate to the extent that they have the same pattern of relations with occupants of other positions. More specifically, two people are structurally equivalent to the extent that they have identical relations with all other individuals in the study population.

Or, as Sailer (1978) puts it, "Two people in the same role are substitutable." Because this dissertation looks at social influence in the context of geography, demography, and ideology, the use of mass survey data in this dissertation is based on the idea that individuals with the same geographical, demographic, or ideological characteristics are structurally equivalent. This means that individuals with the same geographical, demographic, or ideological characteristics have the same patterns of relationships, so that randomly chosen individuals, so long as they have the same characteristics, can be studied to see whether they influence and are influenced by other individuals with similar characteristics.[12]

For example, suppose two individuals reside in the same congressional district. Suppose they know each other and interact on some level with one another, which may result in the mutual influence of each other's political attitudes. In an ideal world of perfect and complete data, both individuals would be included as data points and the amount of social influence can be estimated based on such data. But actual datasets such as the 2004 might include only one or neither individuals. By leveraging the idea of structural equivalence, however, we can assume that even if the randomly selected individuals in the 2004 ANES are virtually strangers, we can still assess whether the randomly selected

[12] The concept of structural equivalence is often contrasted with contagion processes of social influence. For example, Burt (1987) considered the adoption of medical innovation among medical professionals as the result of structural equivalence and as the result of social contagion. The latter process means that common adoption is the result of social influence among interacting agents, while the former means that common adoption is the result of similar circumstances. Thus, the juxtaposition between structural equivalence and contagion is really just a version of Galton's Problem. While use of structural equivalence as a justification for using mass survey data in this dissertation is somewhat unorthodox, I believe that it is in line with the implications of the concept.

individuals influence and influenced by other randomly selected individuals so long as they share their geographical, demographic, ideological characteristics.

Another advantage of using the ANES is that it is a well-known, often-used dataset. Its weaknesses and strengths are known. Therefore, it is well-suited for this dissertation's alternative approach to social influence.

DEPENDENT VARIABLES

The dependent variables are individual-level attitudes toward specific political issues, which are familiar to political scientists. These dependent variables are attitudes toward the following political issues:

1. Whether medical expenses should be covered by government insurance or by private insurance
2. Whether the environment should be protected even if it eliminates jobs or reduces the standard of living
3. Whether the federal government should let everybody get ahead on their own or ensure that every person has a job and a good standard of living
4. Whether the government should provide fewer services to reduce spending or more services even if it means more spending
5. Whether the federal government should make every effort to improve the social and economic position of blacks
6. Whether women should have an equal role with men in business, industry, and government or remain in the home
7. Whether defense spending should be increased or decreased
8. Whether abortions should be more or less restricted

For each of these issues, the survey respondents in the ANES have been asked to place themselves on a seven-point scale. Descriptions and summary statistics for these eight dependent variables can be found in the appendix.

The independent variables used in this dissertation's models also come from the 2004 ANES. They include various socio-economic and demographic variables as well as variables associated with political behavior and attitudes, such as party identification and political knowledge. Subsequent chapters will provide more details on the specific independent variables used in different models.

The main measurement issues in this dissertation have to do with specifying the connectivities among the survey respondents in spatial weights matrices. These specifications have significant consequences for estimating spatial regression models and will be dealt with in the following empirical chapters.

SPATIAL WEIGHTS MATRICES

This subsection discusses methods for locating respondents in three different types of spaces, for calculating distances between respondents in these spaces, and for using these distances to construct spatial weights matrices.

Blau Space

To estimate social distance, it is necessary to define the dimensions of the Blau space. Based on an informal review of political journal articles, I identify race, education, and sex as the most commonly used demographic variables and are therefore well-suited to serve as dimensions of a Blau space. To calculate the distance between individuals i and j in the Blau space, I used Gower's dissimilarity coefficient S_{ij}, which is given by

$$S_{ij} = \frac{\sum_{N} \delta_{ijk} s_{ijk}}{\sum_{N} \delta_{ijk}}$$

Higher values of Gower's S_{ij} indicate greater dissimilarity (or greater social distance), while lower values indicate lesser dissimilarity (or lesser social distance). Using this measure, the elements of the spatial weights matrix **W** are:

$$w_{ij} = \frac{1}{S_{ij}} \qquad \text{for } i \neq j$$

$$w_{ij} = 0 \qquad \text{for } i = j$$

This means that the spatial lag **Wy** can be seen as the weighted average of the respondents' Blau neighbors.

Ideological Space

Using Poole's scaling technique, individuals can be placed in a basic ideological space by using their candidate evaluation scores from public opinion surveys. Poole's technique is based on a singular value decomposition of a numeric matrix with missing data (see Poole, 1998).

By placing each individual on a single basic ideological space, one may compute distances between individuals in the space and assess the ideological similarity between them. The smaller the ideological distance between two respondents, the more similar they are. I expect that respondents are more influenced by those who are more ideologically similar to themselves. This implies that elements of the spatial weights matrix **W** should be:

$$w_{ij} = \frac{1}{d_{ij}} \qquad \text{for } i \neq j$$

$$w_{ij} = 0 \qquad \text{for } i = j$$

where d_{ij} is the Euclidean distance between respondents i and j in the basic ideological space. This means that the spatial lag **Wy** can be seen as the weighted average of the respondents' ideological neighbors.

Geographical Space

In this dissertation, I use the congressional district to represent geographical space. Two respondents are geographical neighbors if they reside in the same congressional district. This means that the elements of the spatial weights matrix **W** are:

$w_{ij} = 1$ if respondents i and j are neighbors

$w_{ij} = 0$ for $i = j$ and if respondents i and j are not neighbors

The spatial lag for the respondents can be seen as the unweighted average of their fellow congressional district dwellers.

INDICATORS OF SPATIAL AUTOCORRELATION

Spatial autocorrelation is the measure of the strength and direction of spatial dependence. Positive spatial autocorrelation indicates that higher values (of the variable of interest) correlate with higher neighboring values and that lower values correlate with lower neighboring values. Negative spatial autocorrelation indicates that higher values correlate with lower neighboring values and that lower values correlate with higher neighboring values. If there is no spatial autocorrelation, there is no association between values and locations.

The most widely-used indicator of spatial autocorrelation is Moran's *I*. By using the (global) Moran's *I* test, one can test the null hypothesis of spatial randomness against the alternative hypothesis of spatial autocorrelation. Moran's *I* is given by:

$$I = \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \tag{3.3}$$

where w_{ij} is spatially-defined weight between spatial units i and j (*e.g.*, individuals, cities, states) and y_i is an observation of the variable of interest (*e.g.,* policy, vote, institution). Moran's I is essentially a ratio of the weighted covariance over the variance of the observations, where w_{ij} are the weights. If the distance between units i and j is large, then w_{ij} will be very small and the cross-product of units i and j's deviations from the mean will be small. On the other hand, if w_{ij} is very large, then the cross-product of units i and j's deviations from the mean will be large. Moran's I ranges from approximately -1 (which indicates negative autocorrelation) to 1 (which indicates positive autocorrelation). If Moran's I is equal to zero, then there is no autocorrelation, implying a random dispersion of values in a space.[13] In a Moran's I test, the null hypothesis is that there is no systematic relationship between values and location, or spatial randomness. The alternative hypothesis is that there is spatial autocorrelation. Positive and significant spatial autocorrelation indicates that there is spatial clustering of the variable of interest.

Moran's I is the most popular indicator of spatial autocorrelation because it is easily applicable and performs well in a variety of situations. For these reasons, it will be used in this dissertation to test for spatial autocorrelation in political attitudes.

But Moran's I is by no means the only indicator of spatial autocorrelation. There are two broad classes of indicators: global and local. *Global indicators of spatial autocorrelation*

[13] The Moran's I test can be conducted under the assumption of normality or by using random permutation. This study uses the latter because it is data-driven and does not make parametric assumptions. The results obtained from Moran's I tests that assume normality, however, are almost identical.

measure the degree of spatial autocorrelation for the entire dataset. These measures are based on the assumption that the same spatial process occurs over the entire space – *i.e.*, no spatial heterogeneity. Global measures include Moran's *I*, Gamma, Geary's *C*, and the Getis-Ord *G* statistics. *Local indicators of spatial autocorrelation* measure the degree of spatial autocorrelation for each observation unit. These measures are based on the assumption of spatial heterogeneity – *i.e.*, the assumption is that different spatial processes may occur over the same space. Local indicators of spatial autocorrelation include local Moran's *I*, local Gamma, local Geary's *C*, and others. Both global and local indicators of spatial autocorrelation are closely related; in general, the sum of a local measure of spatial autocorrelation is proportional to the corresponding global measure (see Anselin, 1995).

MODEL

The spatial lag model is the primary model used in this dissertation. The spatial lag model takes the following form:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
$$\varepsilon \sim N\left(0, \sigma^2\mathbf{I}\right)$$

where $\mathbf{y}$ is the *n* by 1 vector of observations of the dependent variable. $\boldsymbol{\beta}$ is a *k* by 1 vector of parameters associated with independent variables $\mathbf{X}$, which is an *n* by *k* matrix. The error term ε satisfies the usual Gauss Marcov assumptions, and $\mathbf{I}$ is a *n* by *n* identity matrix.

The spatial autoregressive parameter ρ measures the effect of the spatial lag $\mathbf{Wy}$. If $\rho = 0$, then there is no spatial dependence, and the spatial lag model reduces to the classical linear regression model. On the other hand, if the spatial autoregressive parameter ρ is statistically different from zero, there would be evidence of social influence on political attitudes in the social space represented by $\mathbf{W}$.

ESTIMATION METHODS

In general, spatial regression models cannot be estimated using ordinary least squares (OLS). The exception to this is the spatial-x model, which can be estimated using OLS so long as the Gauss-Marcov assumptions hold. For the spatial lag model and FAR model, however, OLS estimation generates biased and inconsistent estimates (Dow, 1982; Anselin, 1988; Le Sage 1998);[14] the presence of y on both sides of the equation gives rise to simultaneity bias. To overcome these problems, maximum likelihood estimation should be used for estimating the parameters of the spatial lag model and the FAR model. Since the FAR model can be considered a special case of the spatial lag model (with no covariates), the discussion focuses on estimating spatial lag models.

To estimate the spatial lag model using maximum likelihood, Anselin (1988) provides the following procedure:

1. Use OLS to estimate $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{OLS} + \boldsymbol{\varepsilon}_{OLS}$
2. Use OLS to estimate $\mathbf{Wy} = \mathbf{X}\boldsymbol{\beta}_{L} + \boldsymbol{\varepsilon}_{L}$
3. Compute the two sets of residuals $\mathbf{e}_{OLS} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{OLS}$ and $\mathbf{e}_{L} = \mathbf{Wy} - \mathbf{X}\hat{\boldsymbol{\beta}}_{L}$
4. Given the two sets of residuals from step 3, find ρ that maximizes the following concentrated likelihood function

$$L_c = C - \frac{n}{2}\ln\frac{1}{n}\left(\mathbf{e}_{OLS} - \rho\mathbf{e}_{L}\right)'\left(\mathbf{e}_{OLS} - \rho\mathbf{e}_{L}\right) + \ln\left|\mathbf{I} - \rho\mathbf{W}\right|$$

5. Using $\hat{\rho}$ from step 4, compute the estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}_{\varepsilon}^{2}$ using

[14] Anselin (1988) showed that the OLS estimate of the spatial autoregressive parameter ρ is inconsistent, since

$$\operatorname{plim} n^{-1}\left(\mathbf{y}'\mathbf{W}'\boldsymbol{\varepsilon}\right) = \operatorname{plim} n^{-1}\boldsymbol{\varepsilon}'\mathbf{W}\left(\mathbf{I} - \rho\mathbf{W}\right)^{-1}\varepsilon$$

which is equal to zero only in the trivial case where $\rho = 0$.

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{OLS} - \rho\hat{\boldsymbol{\beta}}_L$$

and

$$\sigma_{\varepsilon}^{2} = \frac{1}{n}(\mathbf{e}_{OLS} - \rho\mathbf{e}_L)'(\mathbf{e}_{OLS} - \rho\mathbf{e}_L)$$

Using maximum likelihood to estimate the spatial lag model has many advantages. First, maximum likelihood estimates are asymptotically efficient and consistent, so long as certain regularity conditions are met. This means that the spatial autoregressive parameter must be constrained such that $|\rho| < 1$ (Anselin, 1988). Second, maximum likelihood estimation is substantively appropriate. According to Lin et al. (2006), maximum likelihood estimation allows for the endogeneity of *y*, and in doing so, provides an "equilibrium solution" to the dynamic influence process represented in the spatial lag and FAR models. If $\rho \geq 1$, the process can be non-stationary or explosive, which would imply the lack of an equilibrium.

Another popular method for estimating spatial regression model is two-stage least squares. While maximum likelihood estimation is more asymptotically efficient, two-stage least squares can be less computationally intensive. In a two-stage least squares procedure (see Kelejian and Prucha, 1998), one can use $\mathbf{X}$, $\mathbf{WX}$, $\mathbf{W}^2\mathbf{X}$, and other transformations of $\mathbf{X}$ using $\mathbf{W}$ as instruments in a first-stage regression of $y^* = Wy$ to obtain an estimate for the spatial lag term, and then use OLS on the spatial regression model with the estimated spatial lag $\hat{y}^*$.

The decision to use maximum likelihood or two-stage least squares depends on several factors. Lin et al. (2006) found maximum likelihood adequate for estimating spatial lag models with only one spatial lag but turned to two-stage least squares estimation for models with two or more spatial lag terms. Cho (2003) recommended two-stage least squares for large datasets and for data exhibiting a high degree of non-normality. Since each of the models used throughout this dissertation feature only one spatial lag and the dataset of interest features only 1200 observations with low levels of non-normality (in the dependent variables), this dissertation will rely on maximum likelihood estimation.

GALTON'S PROBLEM

If a test for spatial autocorrelation turns out to be statistically significant, the researcher cannot conclude that there is spatial dependence (and therefore social influence) and stop there. Examinations of indicators of spatial autocorrelation are exploratory in nature; they do not account for other factors that may give the *appearance* of spatial dependence. Neighboring values may be similar because neighboring individuals mutually influence each other, or they may be similar because of some coincidence or common predisposition. For example, one might find statistically significant positive spatial autocorrelation in vote choice across counties, which is evidence that adjacent counties tend to vote in similar ways. This relationship between geographical location and vote choice can be explained in different ways. One, there may actually be social influence among neighboring counties, due to the interactions among individuals who live close together in those counties. This would be true spatial dependence. Two, it may be that adjacent counties have faced common external shocks, such as a common media market or environmental issue, that have led them to vote independently in a similar direction. Three, it may be that Republicans tend to live with other Republicans while Democrats tend to live with other Democrats, which means that the similarity in vote choice in neighboring counties is really the result of party affiliations.

Distinguishing between true spatial dependence and the effects of common shocks is known in spatial econometrics as Galton's Problem. This methodological issue is named for Sir Francis Galton, who pointed out in 1889 that inferences based on assumptions of independence can be misleading if there is actually diffusion among observations of observations. The subject of his criticism is an 1889 anthropological study of marriage and descent across cultures. Edward Tylor, who carried out the study, found that certain types of institutions of marriage and descent were associated with each other and concluded that these associations implied a general evolutionary sequence involving institutional shifts from maternal lines to paternal lines. In response, Galton pointed out that similarity among cultures could come from borrowing, common descent, or evolutionary development. Without controlling for these interdependencies, he argued, one cannot make valid inferences across cultures. While it arose out of cross-cultural anthropological research, Galton's problem became a general but serious methodological issue in social science research.

In spatial econometrics, Galton's Problem is the problem of distinguishing between mutual influence and coincidence, between common attributes and interaction, between independent and interdependent processes. By understanding mutual influence as a process, we are saying that individuals become more alike through some process of interaction. This is distinct from saying that individuals are more alike in some ways because they are more alike in other ways. Dow et al.'s (1982) explanation is useful here:

> When interacting units tend to become either more alike – through diffusion, contagion, imitation, assimilation, cooptation, convergent competition or a host of other process – or more dissimilar – through repulsion, divergent competition, differentiation, etc. – as a result of interaction, we have a particular kind of contextual effect resulting from the network of relations between sample unit rather than from their attributes.

Franzese and Hays characterized Galton's Problem as the "great difficulty" of distinguishing *common shocks* from *interdependence*. Common shocks result in "correlated response to correlated unit-level, contextual, or context-conditional factors," while interdependence refers to "contexts in which the outcomes of interest (*i.e.*, dependent variables) in some units of analysis (*e.g.*, countries) directly affect outcomes in others. However, spatial patterns might appear without any direct effect from outcomes in some units to those in others, perhaps via spatial correlation in domestic or exogenous-external conditions affecting units.

In the case of political attitudes, Galton's Problem leads us to the following question: Do some individuals have similar attitudes because they mutually influence each other, or do they have similar attitudes because they were exposed to the same stimulus or were predisposed to have those attitudes in the first place? Answering this question means figuring out whether similarities in political attitudes are due to a process (*i.e.*, mutual influence) or an attribute (*i.e.*, common shock). This can be accomplished to a large degree by using carefully-considered control variables in spatial regression models. Accordingly, each spatial regression model in the following three chapters will feature control variables appropriate to influence in the social space under examination.

HYPOTHESES

The null hypothesis is that there is no neighborhood effect on political attitudes in Blau, ideological, and geographical spaces. The alternative hypothesis is that there is a positive neighborhood effect, meaning the spatial parameters associated with each type of spatial regression model are all greater than 0 and less than 1.

For example, if the estimated spatial autoregressive parameter $\hat{\rho} \neq 0$, then there would be evidence of a neighborhood effect on individual political attitudes.

CONCLUSION

This chapter outlined the next three chapters' empirical research design, including the research question, variables, measurement, models, and hypotheses. The next three chapters apply this design to the study of social influence on political attitudes in geographical, Blau, and ideological spaces.

CHAPTER 5: SOCIAL INFLUENCE AND GEOGRAPHICAL PROXIMITY

In the previous chapters, we saw how social influence can be framed in spatial econometric terms and that spatial regression models offer important advantages over linear regression models and contextual models for modeling individual-level and group-level effects. In this chapter, I apply these ideas by using spatial econometric tools to see whether geographically-proximate individuals mutually influence each other's political attitudes.

Does geographical proximity play a role in social influence? Intuitively, there is reason to suspect that it does. As individuals think about political issues, it is reasonable to suppose that they account for the views of their family, friends, colleagues, and neighbors – people with whom they have frequent contact in their day-to-day lives not only because of their personal relationships, but simply because they reside in the same geographical area.

This chapter has two aims – one substantive, the other methodological. The first aim is to see whether geographically-proximate individuals mutually influence each other's political attitudes. To do this, I first explain how geographical proximity might facilitate social influence. Then I identify a spatial weights matrix that best represents the relationships among individuals in a geographical space. Next, I use the spatial weights matrix and spatial econometric tools to test for spatial dependence and to estimate the neighborhood effect on political attitudes. Based on the results, I assess whether and to what extent geographical proximity abets social influence on political attitudes.

The second aim of this chapter is to illustrate some of the statistical results from chapters 2 and 3. Chapter 2 showed that the contextual model can be considered a special case of the spatial lag model, while chapter 3 showed that the dummy variable regression model (through the origin) is a special case of the first-order spatial autoregressive model. Using congressional districts as mutually disjoint and exhaustive neighborhoods, I compare the results of estimating these models using the same data. This comparison serves to clarify the differences between spatial regression models and more conventional models for studying social influence.

GEOGRAPHY AND SOCIAL INFLUENCE

Researchers have found that geography is important for understanding political behavior and attitudes. Cho (2003) uncovered evidence of a diffusion effect in Asian-American campaign contribution networks within ZIP codes. Lin et al. (2006) found that national identity is contagious within Taiwanese townships. Chen and Rodden (2009) found that the probability that two randomly drawn individuals belong to the same political party is a function of the distance between their residential locations. These and other studies show that geographical proximity matters. They point to a positive relationship between the political attitudes and behaviors of individuals and the political attitudes and behavior of their geographically-proximate peers. But why is there such a relationship? What are the theoretical reasons for political attitudes to be susceptible to social influence due to geography?

One major reason is that geography often constrains human activity and interaction. As individuals carry out their lives, they encounter and interact with other people who often live in the same geographical space. Geographical spaces constrain the pool of potential individuals with whom one interacts (Putnam, 1966; Huckfeldt, 1983a; Huckfeldt, 1986; Huckfeldt, 2009). Given the diversity of individuals and communities and given that individuals sort themselves into communities with similar tastes and values, we would expect heterogeneity across geographical areas, but relative homogeneity within geographical areas. From an individual's perspective, this homogeneity means that there is not an unlimited pool of potential friends and associates in his geographical area with whom to discuss politics or other topics. Huckfeldt (2009) explains,

> Suppose that the supply of discussants with particular sets of preferences is unlimited and that people are infinitely patient in their search for political discussion partners. Under these circumstances, individuals would always get exactly what they want in terms of discussants and communication networks. If the Red Sox fan who is a Democrat wants to find another Democratic Red Sox fan, she or he will do so. And under these circumstances, homophily will be the ever persistent state of affairs within communication networks. The problem is that, first, the supply of Democrats who are also Red Sox fans may be quite limited, at least if you live in New York City. And second, the personal costs of patience are quite high—it may mean social isolation or at least not having anyone with whom to talk politics (or baseball) at lunch.

Because geographical areas constrain associational choices, there is a corresponding constraint on the possible impersonal and interpersonal interactions that can take place within an area. These impersonal and interpersonal interactions provide opportunities for social influence. Thus, geographical areas can give rise to social influence in this indirect way. But how might geographically-constrained interactions directly facilitate social influence?

Geographically-constrained impersonal and interpersonal interactions abet social influence by affecting the marginal utility that an individual derives from holding a certain political attitude. As a result of these interactions, an individual's political attitude can change, resulting in an interdependence of political attitudes in geographical space. These interactions fall into three classes of social influence mechanisms: *coercion*, *learning*, and *emulation*.

Ccoercion refers to the submission of the weaker to the stronger. As Franzese and Hays (2007a) point out, the types of coercion that can lead to interdependence "may be direct or indirect and hard (force) or soft (suasion)." In geographical contexts, individuals might be coerced to conform to certain attitudes or behaviors through *social pressure* or *political mobilization*.

Social pressure refers to the intentional or unintentional pressure that other individuals may exert on individuals to have certain attitudes or engage in certain behaviors. Thus, others' attitudes are influential not only because of their credibility or content, but because of the nature of sociability itself: People are inclined to conform in opinion and action to those around them in order to be liked. Consequently, in the face of disagreement, individuals may change their initial attitudes or behaviors in order to get along. Since physical proximity often enhances social pressure, we should expect that individuals who are geographically close to each other to experience greater social pressure to conform to certain political attitudes.

Political mobilization is an activity designed to make individuals hold certain political attitudes or engage in certain political behaviors. According to Cho (2003), because political campaigns are strategic with their limited resources, political candidates focus on certain media markets but not others: "Because this courting is geographically definable, [campaign] donations may appear to be rolling in in geographic clusters rather than emerging as random, independent events across the United States." Similarly, we might expect that political candidates would target specific electoral groups (such as a congressional district) for persuasive appeals. Because different political candidates specialize in different policies, we would expect them to make different appeals to their specific constituencies. As a result, similar individuals in different geographical areas might end up with different political attitudes because of their exposure to different campaigns.

Learning is a class of social influence mechanisms that involve information exchange. As a mechanism for social influence, learning is based on the presumption that individuals, whatever their initial impulses, are open to persuasion (see Elkins and Simmons, 2005). Whether they like it or not, individuals are exposed to the attitudes or behaviors of other people in their communities and through this exposure they acquire information about political issues. According to Pattie and Johnston (1999, 2000), swing voters who discuss politics with others tend to take on the views of their discussants – a process that they call "conversion by conversation." But information acquisition is not limited to conversations or explicitly political activities. Huckfeldt (1986) points out that one can acquire valuable information about a community's political views simply by observing bumper stickers or yard signs. But because individuals living in geographical contexts are exposed to a limited selection of individuals and a limited selection of viewpoints, there is an information bias with respect to what can be acquired. As a result, what might be learned in one geographical context might be very different from what might be learned in another geographical context.

Emulation refers to the ritualistic adoption of others' attitudes or behaviors. In a geographical context, individuals tend to conform to others' attitudes or behaviors in unconscious and conscious ways. According to the idea of *unconscious emulation*, individuals may gradually form certain attitudes and behaviors because of mere exposure to those attitudes and behaviors. *Conscious emulation* arises from the tendency of individuals to feel rewarded for conforming to the attitudes or behaviors of their peers. Psychological studies have shown that individuals were more likely to give an incorrect answer to a question when other individuals did so, while individuals tended to give correct answers when there was no similar pressure to conform (Asch, 1951). According to threshold models of behavior (Schelling, 1978; Granovetter, 1978), adopting the dominant attitude or engaging in the dominant behavior enhances its legitimacy. As the number or proportion of those who hold certain attitudes or engage in certain behaviors increase, the degree of legitimacy of those attitudes or behaviors increases in the eyes of the individual.

Because of coercion, learning, and emulation, we would expect individuals who are geographically-proximate to have more similar political attitudes than individuals who are far apart. As Huckfeldt (1986) observed, "[I]ndividual behavior tends to move in the direction of the surrounding population's social makeup, even when individual characteristics are taken into account."

Furthermore, because individuals are members of their community, the social influence arising from coercion, learning, and emulation cannot be considered a one-way relationship. It would be illogical to suppose that an individual is affected by others in his geographical contexts, but that others are not affected by what the individual does – after all, communities consist of individuals. Once an individual adopts an attitude or engages in a behavior, the process of social influence does not stop there; that individual also affects others in the same geographical context. Individuals may be softly coerced into adopting a certain political attitude, but once they do so, their adoption strengthens the soft coercion that other individuals face. After acquiring (biased) political information, individuals may in turn pass the information onto others and affect their political attitudes. Individuals who consciously or unconsciously conform to a certain political attitude become part of the bandwagon, thereby adding to the utility of conformity and/or enhancing the exposure to that political attitude.

At the same time, we cannot suppose that individuals mutually influence each other *ad infinitum*, for that would require the assumption that individuals mindlessly and selflessly follow the attitudes of others. Such an assumption would do great violence to the idea of free will and the very possibility of politics. Political attitudes are meaningful so long as they are tied to actual values and ideas, which are ultimately rooted in individual rumination in light of, but not based solely on, prevailing attitudes and behaviors. Social influence is naturally limited by individual thought; and though this limit might differ from individual to individual, this means that the process of social influence does not alter political attitudes forever, but reaches some kind of equilibrium in which individuals reconcile their personal and social inclinations and have no further incentive to change.

All this suggests that while the social influence among geographically proximity individuals must be considered as *mutual* influence, this influence is naturally limited. As we will see, the spatial lag model captures this idea of mutual influence in equilibrium, while the linear regression and contextual models do not.

METHODOLOGICAL APPROACH

I operationalize the idea of geographical proximity by mapping survey respondents from the 2004 ANES onto their respective congressional districts and using this information to construct several spatial weights matrices. Using these matrices, I test for spatial autocorrelation to see whether there is a relationship between the respondents' geographical locations and their attitudes toward eight political issues. If there is evidence of spatial autocorrelation for a particular political attitude, I see whether the spatial autocorrelation holds after controlling for individual-level explanatory and control variables.

Dependent and Independent Variables

The focus of this chapter is the analysis of eight individual-level political attitudes from the 2004 ANES. These dependent variables are respondents' self-placements on the following seven-point political issue scales:

1. Whether medical expenses should be covered by government insurance or by private insurance
2. Whether the environment should be protected even if it eliminates jobs or reduces the standard of living
3. Whether the federal government should let everybody get ahead on their own or ensure that every person has a job and a good standard of living
4. Whether the government should provide fewer services to reduce spending or more services even if it means more spending
5. Whether the federal government should make every effort to improve the social and economic position of blacks
6. Whether women should have an equal role with men in business, industry, and government or remain in the home
7. Whether defense spending should be increased or decreased
8. Whether abortions should be more or less restricted

Detailed descriptions and summary statistics for these eight dependent variables can be found in the appendix.

Independent variables include control variables and explanatory variables that may be relevant to political attitude formation. These include the following individual-level socio-demographic (control) variables:

1. Age (in years)
2. Number of children

3. Whether the respondent is female
4. Whether the respondent lives in a rural environment
5. Whether the respondent is married
6. Whether the respondent belongs to the military
7. Social class (self-placed scale, where higher values correspond to higher social classes)

Explanatory variables include the following individual-level variables:

1. Religiosity (whether the respondent attends church)
2. Intelligence (based on the interviewer's assessment on a 5-point scale, where higher values correspond to greater intelligence)
3. Household income
4. Years of education
5. Level of interest in political campaigns (based on a 3-point scale, where higher values correspond to greater interest)
6. Level of political knowledge (based on the number of correct responses to three knowledge questions)
7. Party identification (based on a 7-point self-placement scale, where higher values correspond to stronger identification with the Republican Party and lower values correspond to stronger identification with the Democrat Party)
8. Political ideology (based on a 7-point self-placement scale, where higher values correspond to greater conservatism and lower values correspond to greater liberalism)
9. The importance of the political issue (on a 5-point scale, where higher values correspond with greater importance to the respondent)

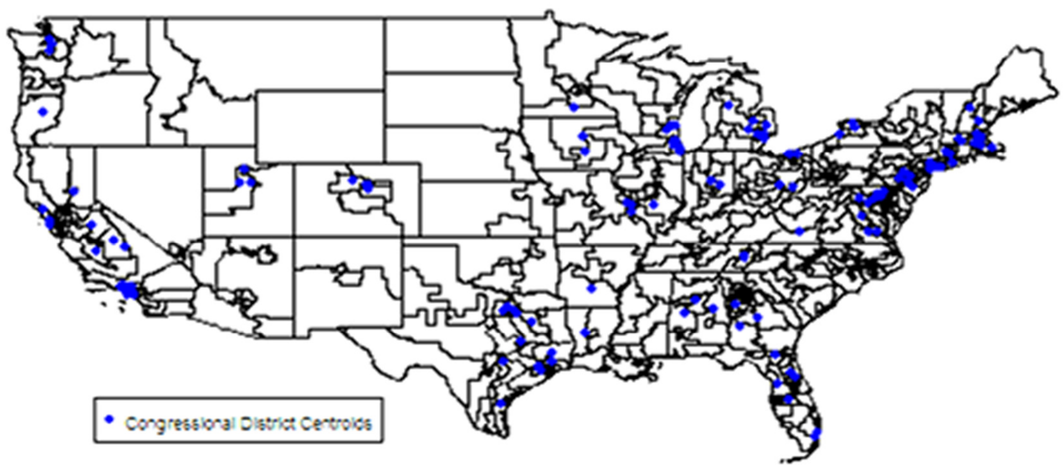
Operationalizing Geographical Proximity

The unit of analysis is the individual, which comes with some challenges for the spatial econometric analysis in this chapter. Since I am using the 2004 ANES, there are no data on the exact locations of the survey respondents, such as the latitude and longitude of their homes, or even the towns in which they live. However, the 2004 ANES does identify the congressional districts to which they belong, which can provide a rough measure of where the respondents are positioned geographically.

The 1212 survey respondents for the 2004 ANES came from 140 congressional districts and 29 states. To map the survey respondents, I looked up the coordinates of each congressional district's centroid and assigned each set of coordinates to the corresponding respondents. Figure 1 is a map of congressional districts, where the dots mark each district's centroid.

After mapping all the respondents to their congressional districts, I calculated the distance between the geographical centers of each pair of congressional districts. These distances can be used to construct spatial weights matrices.

Figure 5.1. Geographical Locations of 2004 ANES Survey Respondents



Based on 1212 respondents from the 2004 ANES

The Spatial Weights Matrix

In a spatial regression analysis, the specification of the spatial weights matrix is considered known. I will consider five possible definitions of geographical proximity and five corresponding spatial weights matrices and show why a spatial weights matrix based on common congressional district is most appropriate for the analysis in this chapter.

The first plausible definition of geographical connectivity between individuals in the dataset is based on common congressional district. The mutual influence of political attitudes among individuals may be more likely to occur between individuals who reside in the same congressional district. Under this definition of connectivity, the spatial weights matrix is a binary connectivity matrix where each element w_{ij} is equal to 1 if individuals i and j reside in the same congressional district. An individual's neighbors are individuals who live in the same congressional district, and each congressional district is a disjoint neighborhood.

The spatial lag for each individual can be seen as the unweighted average of fellow congressional district denizens.

A second plausible definition of connectivity is based on neighboring congressional districts. Individuals may be more likely to influence each others' political attitudes if they reside in the same congressional district or if they reside in adjacent congressional districts. This definition of connectivity yields two spatial weights matrices:

- *K-nearest neighbor*. An individual's neighbors are those who live in the *k* congressional districts that are closest to the individual's congressional district, where *k* is a positive integer. "Closeness" is based on the distances between the congressional districts' centroids.

- *Distance band*. An individual's neighbors are those who live either in the same congressional district or in congressional districts within a 400-kilometer distance band of the individual's congressional district's centroid.

For both of these spatial weights matrices, the spatial lag for each individual can be seen as the average of fellow and nearby congressional district denizens.

The problem with the *k*-nearest neighbor and distance band matrices is that they are not feasible. The 2004 ANES may be based on a random sample of individuals and represent Americans, but it lacks representation from all the congressional districts – the dataset includes only 140. Therefore, identifying the *k* nearest neighbors is fine in theory, but given this dataset, the available data points may exist in congressional districts that are too far away to be plausibly influential; for example, there is only one respondent from a congressional district in Oregon, which means that the closest neighbors must come from congressional districts in Washington and California. As for the distance band spatial weights matrix, the 400-kilometer distance threshold for influence creates problems because respondents in smaller congressional districts would have a lot of neighbors but respondents in large congressional districts would have very few neighbors. Given that congressional district sizes are based on population, using the distance band spatial weights matrix may produce misleading results.

A third plausible definition of connectivity is based on inverse distance. This is based on the idea that individuals are more likely to influence each other's political attitudes when they reside closer together and are less likely to influence each other when they reside farther apart. Under this definition of connectivity, there are two spatial weights matrices for consideration:

- *Inverse distance*. All individuals are neighbors of all the other individuals, but individuals are more "neighborly" if their congressional districts are the same or if their congressional districts are closer together. In contrast, individuals are less "neighborly" if their congressional districts are farther apart. "Closeness" is based on the distances between the congressional districts' centroids; greater distances indicate less neighborliness.

- *Inverse distance squared*. All individuals are neighbors of all the other individuals, but individuals are more "neighborly" if their congressional districts are the same or if their congressional districts are closer together. In contrast, individuals are less "neighborly" if their congressional districts are farther apart. "Closeness" is based on the squared distances between the congressional districts' centroids; greater squared distances indicate less neighborliness.

For both of these spatial weights matrices, the spatial lag for each individual can be seen as the weighted average of fellow and nearby congressional district denizens. The weights are the inverse distance and inverse distance squared.

Using a spatial weights matrix based on inverse distance or inverse distance squared is not feasible when using the 2004 ANES dataset because the distance measures are too rough. Since the location of each respondent is approximated with the geographical center of his congressional district, the inverse distance or inverse distance squared between two respondents who live in the same congressional district would be undefined, which means that the respondents who should count the most are not counted at all. Even so, this problem is not insurmountable; if the distances between respondents in the same congressional districts were gently massaged so that they are very small numbers rather than zero, respondents who live in the same congressional district could still count for a great deal. But there still remains the problem of respondents who live very far away. Though the spatial weights matrix and corresponding spatial lag term would show that their influence would be very close to zero, in reality, it is very implausible that there would be any sort of interaction at all between respondents who live two or three congressional districts over, much less in another state.

Table 5.1 summarizes the possible spatial weights matrices.

Table 5.1. Spatial Weights Matrices for Geographical Proximity

TYPE OF CONNECTIVITY	DEFINITION OF INDIVIDUAL I'S NEIGHBORHOOD	DEFINITION OF SPATIAL WEIGHT W_{IJ}
Same district	Those who live in the same congressional district	$w_{ij} = 1$ if individuals i and j live in the same congressional district $w_{ij} = 0$ otherwise or $i = j$
K-nearest neighbor	The *k* nearest individuals, where distance is based on the location of their congressional district	$w_{ij} = 1$ if individual j is one of individual i's k-nearest neighbors $w_{ij} = 0$ otherwise or $i = j$
Distance band	Those who live in congressional districts within a 400 km distance band	$w_{ij} = 1$ if individual i's and j's congressional districts are within 400 km $w_{ij} = 0$ otherwise or $i = j$
Inverse distance	All other individuals	$w_{ij} = d_{ij}$ where d_{ij} = inverse distance (km) between individual i's and j's congressional districts $w_{ij} = 0$ if $i = j$
Inverse distance squared	All other individuals	$w_{ij} = (d_{ij})^2$ where d_{ij} = inverse distance (km) between individual i's and j's congressional districts $w_{ij} = 0$ if $i = j$

Because of data limitations and the substantive reasons discussed above, I will use the same-district spatial weights matrix. This means that those who live in the same congressional district will be considered as "neighbors" and that each congressional district will be considered as a disjoint "neighborhood."

Models and Methods

To test for spatial autocorrelation, I calculate Moran's *I* (using the spatial weights matrices described above) to see whether there is a relationship between the respondents' geographical locations and their attitudes toward eight political issues. If there is spatial autocorrelation, then one should expect a statistically significant association between values in a given location with values in neighboring locations. A statistically significant result for a Moran's *I* test indicates that there is evidence of spatial dependence for a political attitude.

If there is spatial autocorrelation for a particular political attitude, I then control for several individual-level explanatory and control variables in a spatial lag model to see whether the spatial autocorrelation still holds. Recall that the spatial lag model takes the following form:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$
$$\varepsilon \sim N\left(0, \sigma^2\mathbf{I}\right)$$

where $\boldsymbol{\beta}$ is a *k* by 1 vector of parameters associated with independent variables **X**, which is an *n* by *k* matrix of personal characteristics that may be relevant to an individual's political attitude, such as intelligence, partisanship, marital status, interest in politics, and religiosity. By assumption, the error term ε is normally distributed with a constant variance, and **I** is a *n* by *n* identity matrix. **Wy** is the spatial lag of **y** and is the average attitude of respondent *i*'s neighbors. ρ is the spatial autoregressive parameter; if $\rho = 0$, then there is no spatial dependence, and the spatial lag model reduces to the classical linear regression model. By estimating ρ, we can see whether there is evidence that geographically-proximate individuals mutually influence each other's political attitudes.

Because the dependent variable appears on both sides of the equation, the spatial lag model cannot be estimated using ordinary least squares (OLS). The OLS estimates will be biased and inconsistent due to simultaneity bias. In contrast, maximum likelihood (ML) estimation can account for the endogeneity of the spatial lag **Wy** and is consistent and asymptotically efficient if the model is correctly specified (Anselin, 1988; Franzese and Hays, 2007b). Because of these desirable statistical and substantive properties, I use ML estimation for the spatial lag models.

For comparison, I also estimate (using OLS) baseline linear regression models with no geographical variables and contextual models with a contextual variable indicating the mean political attitude of each respondent's congressional district. I also estimate a linear regression model with only the congressional district wide average and a first-order spatial autoregressive model. This will illustrate the difference between linear regression models and spatial regression models discussed in chapter 3.

HYPOTHESES

Using a (global) Moran's *I* test, I test the null hypothesis of spatial randomness against the alternative hypothesis of spatial autocorrelation. If there is no spatial autocorrelation, there would be no association between values and locations. Because I believe that geographically-proximate individuals mutually influence each other's political attitudes, I expect to find evidence of a systematic relationship between values and geographic locations, or positive spatial autocorrelation.

For the spatial lag models, I test whether the spatial autoregressive parameter ρ is statistically significant after accounting for explanatory and control variables. The null hypothesis is that $\rho = 0$, and the alternative hypothesis is that $\rho \neq 0$. I expect to find evidence of a neighborhood effect, but I remain agnostic as to whether this effect is positive or negative. In light of the theoretical discussion, I expect social influence on political attitudes to reach an equilibrium, which means that there is a bound on ρ such that $|\rho| < 1$.

ASSESSING SPATIAL DEPENDENCE: RESULTS

The Moran's I tests showed that there is evidence of spatial dependence in political attitudes toward three out of eight political issues (see Table 5.2). There is statistically significant (at the 0.10 level or better) positive spatial autocorrelation for the issues of government versus private medical insurance ($I = 0.040$), the government's role in securing jobs and a good standard of living ($I = 0.090$), and access to abortion ($I = 0.089$). Substantively, these results mean that individuals who live in the same congressional district have more similar political attitudes compared with individuals who live in different congressional districts.

Table 5.2. Results of Moran's I Tests: Same District Spatial Weights Matrix

POLITICAL ATTITUDE	Moran's *i*	SE	P-VALUE
Government versus private medical insurance	0.040*	1.7	0.099
Environment versus jobs tradeoff	0.011	0.48	0.63
Government's role in securing jobs and a good standard of living	0.090***	3.6	0.000
Government spending and services	0.03	1.2	0.21
Government assistance to blacks	0.024	1.0	0.31
Role of women	0.028	1.2	0.21
Defense spending	0.037	1.5	0.12
Access to abortion	0.089***	3.8	0.00

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

COMPARISON OF LINEAR REGRESSION AND SPATIAL REGRESSION

Chapter 3 showed that the dummy variable regression model (through the origin) is a special case of the first-order spatial autoregressive (FAR) model when the spatial autoregressive parameter ρ is constrained so that $\rho = 1$ and the neighborhoods are mutually exclusive, exhaustive, and discrete. The constraint on the spatial autoregressive parameter requires the assumption that individuals in disjoint neighborhoods are completely dominated by their neighbors' attitudes. On the other hand, the FAR model allows for the estimation of ρ, which measures the effect of the neighborhood effect (namely, the size of the social influence) without assuming that individuals who belong to a certain neighborhood automatically conform to the mean attitude of that neighborhood. This section demonstrates that the presence or absence of a constraint on ρ has important consequences for statistical inference.

Using congressional districts as disjoint neighborhoods, I estimated two types of naïve models for the attitudes toward the issues of government versus private medical insurance, the government's role in securing jobs and a good standard of living, and access to abortion. The first type of model is a regression of the individuals' attitudes on the mean attitudes of the individuals' corresponding congressional district, which is an abbreviated version of the dummy variable regression (through the origin), which I will call abbreviated dummy model.[15] The second model is a FAR model, which is a regression of the individuals' attitudes on the spatial lags. These models do not contain other explanatory or control variables.

The results are presented in Tables 5.3a-5.3c. For the first type of model, we can see that the estimated effect of the mean congressional district attitude is exactly 1 for each of the three political attitudes. This is what we would expect based on the theoretical results. In contrast, the estimated neighborhood effects for the spatial lag models are quite different from $\hat{\rho} = 1$. For the issue of government versus private medical insurance, the

[15] Running dummy variable regression models in this application would require 140 dummy variables, which take up too much space without providing much more information.

neighborhood effect is 0. $\hat{\rho} = 0.053$, which is not statistically significant. For the issue of the government's role in securing jobs and a good standard of living, $\hat{\rho} = 0.098$, which is statistically significant at the 0.01 level. Finally, the neighborhood effect for the issue of access to abortion is $\hat{\rho} = 0.062$, which is also statistically significant at the 0.01 level.

These results show that even when a neighborhood effect is weak (i.e., ρ is not statistically significant), the abbreviated dummy variable model can produce estimates showing otherwise.

Table 5.3a. Attitude Models: Government versus Private Medical Insurance

	ABBREVIATED DUMMY VARIABLE MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	0.0 (0.27)	3.5*** (0.11)
Average attitude of fellow district dwellers	1.0*** (0.072)	-
Spatial autoregressive parameter (ρ)	-	0.053 (0.033)
AIC	*3226*	*3394*
Log likelihood		*-1694*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Table 5.3b. Attitude Models: Government's Role in Securing Jobs and a Good Standard of Living

	ABBREVIATED DUMMARY VARIABLE MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	0.0 (0.27)	4.0*** (0.11)
Average attitude of fellow district dwellers	1.0*** (0.061)	-
Spatial autoregressive parameter (ρ)	-	0.098*** (0.028)
AIC	*3091*	*3308*
Log likelihood		*-1651*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Table 5.3c. Attitude Models: Access to Abortion

	ABBREVIATED DUMMY VARIABLE MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	0.0 (0.16)	2.7*** (0.061)
Average attitude of fellow district dwellers	1.0*** (0.055)	-
Spatial autoregressive parameter (ρ)	-	0.062*** (0.025)
AIC	*2353*	*2625*
Log likelihood		*-1309*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

In the next section, I investigate whether the spatial dependence detected in the Moran's *I* tests and naïve models holds after controlling for political knowledge, partisanship, and other individual-level independent variables relevant to political attitudes. In light of

Galton's Problem (see chapter 4), it is important to account for factors other than mutual influence that might give rise to spatial dependence. The positive relationship between geographical location and attitudes toward the issue of government versus private medical insurance may be due to common fate rather than mutual influence. For example, individuals in a particular congressional district may favor government medical insurance over private medical insurance not because they mutually influence each other's attitudes, but because they all happen to belong to the military. Galton's problem can be addressed, to some extent, by controlling for military membership and other variables that might give rise to the spatial clustering of political attitudes.

To conserve space, I present results only for the three dependent variables that returned statistically significant positive autocorrelation in Moran's *I* tests, which are government versus private medical insurance, the government's role in securing jobs and a good standard of living, and access to abortion.[16]

ESTIMATING THE SPATIAL LAG MODEL: RESULTS AND DISCUSSION

I used maximum likelihood to estimate the parameters of the spatial lag models for the issues of government versus private medical insurance, the government's role in securing jobs and a good standard of living, and access to abortion. The results are presented in Tables 5.4a-5.4c, alongside estimates for a baseline linear regression model (with no geographical variables) and a contextual model (with the mean political attitude of each respondent's congressional district as an independent variable) for comparison.

The results show that there is little evidence of a neighborhood effect in political attitudes.

[16] The estimates for the spatial lag models for the dependent variables without statistically significant spatial autocorrelation were not very different from those of the classical linear models. This is unsurprising because in the absence of spatial dependence, the spatial lag model reduces to the classical linear model.

1. For the issue of government versus private medical insurance, the neighborhood effect is $\rho = 0.05$, which is almost statistically significant (p-value = 0.11).

2. For the issue of the government's role in securing jobs and a good standard of living, the neighborhood effect is $\rho = 0.098$, which is statistically significant at the 0.05 level (p-value = 0.00055).

3. For the issue of access to abortion, the neighborhood effect is $\rho = 0.062$, which is statistically significant at the 0.05 level (p-value = 0.015).

The null hypothesis for each spatial lag model is that there is no neighborhood effect ($\rho = 0$). Based on these results, I reject the null hypothesis for the issue of access to abortion and the government's role in securing jobs and a good standard of living, and I accept the null hypothesis for the issue of government versus private medical insurance.

Table 5.4a. Attitude Models: Government versus Private Medical Insurance

	BASELINE LINEAR MODEL	CONTEXTUAL MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	-0.013 (0.46)	-2.3*** (0.48)	-0.048 (0.46)
Age	0.0050 (0.0045)	0.0030 (0.0042)	0.0049 (0.0045)
Children	-0.0065 (0.069)	-0.036 (0.064)	-0.0075 (0.068)
Female	-0.092 (0.13)	-0.028 (0.12)	-0.087 (0.13)
Rural	0.38** (0.16)	0.28* (0.15)	0.38** (0.15)
Married	-0.043 (0.12)	-0.04 (0.11)	-0.045 (0.12)
Military	-0.24 (0.19)	-0.12 (0.17)	-0.24 (0.18)
Attends church	-0.0021 (0.13)	-0.086 (0.12)	-0.0064 (0.13)
Social class	0.12*** (0.039)	0.094*** (0.036)	0.12*** (0.038)
Intelligence	0.081 (0.093)	0.12 (0.087)	0.085 (0.092)
Household income	0.036 (0.012)	0.030*** (0.011)	0.036*** (0.012)
Education	0.026 (0.047)	0.0070 (0.044)	0.025 (0.047)
Interest in campaigns	-0.038 (0.10)	-0.0090 (0.094)	-0.038 (0.099)
Political knowledge	-0.042 (0.055)	-0.047 (0.051)	-0.042 (0.054)
Party ID	0.089*** (0.034)	0.069** (0.032)	0.088*** (0.034)
Political ideology	0.45*** (0.056)	0.069*** (0.032)	0.45*** (0.056)
Issue importance	-0.0065 (0.062)	-0.00058 (0.058)	-0.008 (0.062)
Average attitude of fellow district dwellers	-	0.78*** (0.069)	-

Table 5.4a. Attitude Models: Government versus Private Medical Insurance

(continued)

Spatial autoregressive parameter (ρ)	-	-	0.019 (0.030)
AIC	*3214*	*3096*	*3216*
Log likelihood	*-1589*	*-1529*	*-1589*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Table 5.4b. Attitude Models: Government's Role in Securing Jobs and a Good Standard of Living

	BASELINE LINEAR MODEL	CONTEXTUAL MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	2.1*** (0.45)	-0.81* (0.47)	2.0*** (0.45)
Age	0.0084* (0.0044)	0.0053 (0.0040)	0.0083* (0.0043)
Children	-0.091 (0.064)	-0.068 (0.058)	-0.093 (0.063)
Female	-0.080 (0.12)	-0.030 (0.11)	-0.064 (0.12)
Rural	0.17 (0.15)	0.098 (0.14)	0.17 (0.15)
Married	-0.19* (0.11)	-0.189* (0.10)	-0.20* (0.11)
Military	0.059 (0.18)	0.14 (0.16)	0.054 (0.17)
Attends church	-0.23* (0.12)	-0.21* (0.11)	-0.25** (0.12)
Social class	0.12*** (0.036)	0.097*** (0.033)	0.12*** (0.036)
Intelligence	0.070 (0.086)	0.11 (0.079)	0.085 (0.085)
Household income	0.030*** (0.011)	0.015 (0.010)	0.029*** (0.011)
Education	-0.094** (0.045)	-0.090** (0.041)	-0.099** (0.045)
Interest in campaigns	0.086 (0.096)	0.039 (0.088)	0.084 (0.095)
Political knowledge	0.052 (0.052)	0.022 (0.048)	0.050 (0.052)
Party ID	0.23*** (0.033)	0.16*** (0.031)	0.22*** (0.032)
Political ideology	0.32*** (0.053)	0.31*** (0.049)	0.32*** (0.052)
Issue importance	-0.28*** (0.065)	-0.17*** (0.060)	-0.27*** (0.064)

Table 5.4b. Attitude Models: Government's Role in Securing Jobs and a Good Standard of Living

(continued)			
Average attitude of fellow district dwellers	-	0.74*** (0.060)	-
Spatial autoregressive parameter (ρ)	-	-	0.058** (0.025)
AIC	*3077*	*2935*	*3074*
Log likelihood	*-1521*	*-1449*	*-1518*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Table 5.4c. Attitude Models: Access to Abortion

	BASELINE LINEAR REGRESSION MODEL	CONTEXTUAL MODEL	SPATIAL LAG MODEL
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	3.8*** (0.24)	1.34*** (0.28)	3.7*** (0.241)
Age	-0.0037 (0.0025)	-0.0036 (0.0022)	-0.0038 (0.0024)
Children	-0.078** (0.037)	-0.073** (0.033)	-0.081** (0.036)
Female	0.21*** (0.072)	0.21** (0.065)	0.22*** (0.071)
Rural	-0.19** (0.084)	-0.096 (0.078)	-0.17** (0.083)
Married	-0.089 (0.063)	-0.045 (0.056)	-0.091 (0.062)
Military	-0.060 (0.10)	-0.083 (0.090)	-0.065 (0.099)
Attends church	-0.44*** (0.071)	-0.31*** (0.065)	-0.44*** (0.070)
Social class	0.029* (0.021)	0.006 (0.019)	0.027* (0.021)
Intelligence	0.041 (0.050)	0.088** (0.044)	0.049 (0.049)
Household income	0.033*** (0.0064)	0.021*** (0.058)	0.032*** (0.0063)
Education	0.043* (0.026)	0.028 (0.023)	0.041 (0.025)
Interest in campaigns	0.033 (0.054)	0.028 (0.049)	0.031 (0.053)
Political knowledge	0.10*** (0.030)	0.070*** (0.027)	0.10*** (0.029)
Party ID	-0.064*** (0.018)	-0.054*** (0.017)	-0.066*** (0.018)
Political ideology	-0.16*** (0.030)	-0.11*** (0.027)	-0.15*** (0.030)
Issue importance	-0.18*** (0.033)	-0.18*** (0.030)	-0.18*** (0.032)
Average attitude of fellow district dwellers	-	0.80*** (0.056)	-

Table 5.4c. Attitude Models: Access to Abortion

(continued)

Spatial autoregressive parameter (ρ)	-	-	0.051** (0.022)
AIC	*2396*	*2213*	*2393*
Log likelihood	*-1180*	*-1088*	*-1177*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Compared with the baseline regression model, the spatial lag model fits the data better for the issues of access to abortion and the government's role in securing jobs and a good standard of living. All things being equal, respondents who live in the same congressional district mutually influence each other's attitudes regarding the issues of government versus private medical insurance and access to abortion.

Of the three types of models, the contextual models consistently have statistically significant (at the 0.05 level or better) contextual effects and the best goodness-of-fit numbers. But this is misleading. Since the contextual variable is the mean political attitude of each individual's congressional district and the mean includes the individual's attitude, there is a simultaneity bias that produces inconsistent and biased estimates. Accordingly, we can see that the estimated effects of the independent variables for the contextual models are very different from the corresponding effects of independent variables in the baseline linear regression models and spatial lag models. Furthermore, we can see from the contextual model for the issue of government versus private medical insurance that there is a positive contextual effect ($b = 0.78$) that is very statistically significant (at the 0.01 level); in contrast, the spatial lag model shows that the neighborhood effect is not important at all. This is an example of how contextual models can overstate the neighborhood effects.

Direct Effects, Indirect Effects, and Total Effects

In the spatial lag model, the presence of the dependent variable *y* on either side of the equation means that the fitted coefficients must be interpreted in light of indirect, direct, and total effects. This is because the total effects of the independent variables requires accounting for the spatial multiplier $(\mathbf{I} - \rho\mathbf{W})^{-1}$, since the expected value of *y* is

$$E(y) = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta$$

If there is no neighborhood effect, then $\rho = 0$ and the expected value of *y* reduces to $\mathbf{X}\beta$.

Tables 5.5a and 5.5b contrast the effects found by using the linear regression model with the total effects found by using the spatial lag model for the issues of access to abortion and the government's role in securing jobs and a good standard of living. The presence of a statistically significant neighborhood effect means that the total effect of each independent variable is actually greater than the effect given by the linear regression model. Note that while the marginal effects (1) from the linear regression model are very similar to the direct effects (2) from the spatial lag model, they are quite different from the total effects (4). That is because the total effect of an independent variable for a spatial lag model is the sum of the direct and indirect effects.

The distinction between indirect, direct, and total effects in the spatial lag models shows that endogeneity has consequences beyond the political attitudes of the respondents. Not only do respondents' political attitudes mutually influence each other, their individual characteristics have indirect influences as well. For example, a respondent who identifies as a conservative is more likely to favor restricting access to abortion; the direct effect of being a conservative is -0.15. But since his political attitude also affects his geographical neighbors, there is an indirect effect of his conservatism on their attitudes and an indirect effect of their conservatism toward his attitude as well. In this way, the spatial lag model represents a parsimonious but substantively interesting mechanism of social influence that brings together individual and neighborhood.

Table 5.5a. Comparison of Effects for the Issue of the Government's Role in Securing Jobs and a Good Standard of Living

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL		
	(1)	(2)	(3)	(4)
	b	b (Direct Effect)	Indirect Effect	Total Effect
Age	0.0084	0.0083	0.00034	0.0087
Children	-0.091	-0.093	-0.0038	-0.097
Female	-0.080	-0.064	-0.0026	-0.066
Rural	0.17	0.17	0.0070	0.18
Married	-0.19	-0.19721	-0.00804	-0.20525
Military	0.059	0.054	0.0022	0.056
Attends church	-0.23	-0.25	-0.010	-0.26
Social class	0.12	0.12	0.0050	0.13
Intelligence	0.070	0.085	0.0035	0.088
Household income	0.030	0.029	0.0012	0.030
Education	-0.094	-0.099	-0.0041	-0.10
Interest in campaigns	0.086	0.084	0.0034	0.087
Political knowledge	0.052	0.050	0.0020	0.052
Party ID	0.23	0.22	0.0092	0.23
Political ideology	0.32	0.32	0.013	0.34
Issue importance	-0.28	-0.27	-0.011	-0.29

Table 5.5b. Comparison of Effects for the Issue of Access to Abortion

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL		
	(1)	(2)	(3)	(4)
	b	b (Direct Effect)	Indirect Effect	Total Effect
Age	-0.0037	-0.0038	-0.00014	-0.0039
Children	-0.078	-0.081	-0.0029	-0.083
Female	0.21	0.218	0.0078	0.23
Rural	-0.19	-0.17	-0.0062	-0.18
Married	-0.089	-0.091	-0.0032	-0.094
Military	-0.060	-0.065	-0.0023	-0.067
Attends church	-0.44	-0.44	-0.016	-0.46
Social class	0.029	0.027	0.00097	0.028
Intelligence	0.041	0.049	0.0018	0.051
Household income	0.033	0.032	0.0012	0.034
Education	0.043	0.041	0.0015	0.042
Interest in campaigns	0.033	0.031	0.0011	0.032
Political knowledge	0.10	0.10	0.0036	0.10
Party ID	-0.064	-0.066	-0.0024	-0.068
Political ideology	-0.16	-0.15	-0.0055	-0.16
Issue importance	-0.18	-0.18	-0.0065	-0.19

CONCLUSIONS

This chapter illustrates a spatial econometric approach to the study of social influence on political attitudes in a geographical context. This approach depends on the specification of a spatial weights matrix to represent the structure of mutual influence among geographical neighbors. In this chapter's analysis, geography (operationalized as congressional districts) is the space in which mutual influence might affect the political attitudes of individuals. Each congressional district represents a disjoint neighborhood in which mutual influence can take place. The motivating theory is that geographical proximity supports opportunities for coercion, learning, and emulation, which leads to a similarity of political attitudes within congressional districts.

To examine whether there is evidence that geographically-proximate individuals mutually influence each other's political attitudes, I first used Moran's *I* to test for spatial autocorrelation, thereby uncovering evidence of spatial dependence. Next, I compared the estimates of two naïve models of social influence and found that the abbrebriated dummy variable model tended to overstate the extent of social influence. Finally, I used spatial lag models to estimate neighborhood effects. The results showed that geographically-proximate individuals mutually influence each other's political attitudes regarding two issues: access to abortion and the government's role in securing jobs and a good standard of living. The statistically significant estimates of the spatial autoregressive parameter ρ mean that individual attitudes toward these two issues are subject to mutual influence within congressional districts. Ignoring this effect means implicitly assuming that individuals take on political attitudes without regard to the political attitudes of other geographically proximate individuals.

At this point, it is uncertain why there is a neighborhood effect for only two out of eight political attitudes from the 2004 ANES. These two issues may have been particularly salient during the time of the 2004 ANES data collection, and as a result, individuals were more likely to pay attention to others' attitudes toward them. Or perhaps these two issues may be particularly polarizing issues that represent the ideological leanings of congressional districts, which, we must remember, are political boundaries drawn up to reflect particular ideological leanings. At any rate, rather than speculate regarding the reasons, this chapter goes only as far as to illustrate the application of spatial regression models to the question of social influence on political attitudes. Spatial regression models can identify neighborhood effects for some political attitudes but not for others; further research is needed to determine why some issues are more susceptible to mutual influence and others are not.

The next chapter features another application of this dissertation's spatial econometric approach: an analysis of social influence on political attitudes in Blau space.

CHAPTER 6: SOCIAL INFLUENCE AND SOCIAL SIMILARITY

While the previous chapter analyzed the role of geographical proximity in social influence, this chapter features a spatial econometric analysis of political attitudes in one type of non-geographical space: Blau space. Like geographical proximity, social similarity – understood in terms of proximity in Blau space – might also provide opportunities for individuals to mutually influence each other's political attitudes. As we saw in chapter 2, the idea of space need not be restricted to geography – individuals can mutually influence each other in non-geographical contexts as well.

As individuals make up their minds and feelings about political issues, they often account for the views of their family, friends, colleagues, and neighbors, and these people tend to share similar demographic characteristics, such as race, religion, and education. These characteristics can be viewed as the dimensions of a Blau space. In Blau space, each individual has a position described by *k* coordinates, and distances in the coordinate system define the relationships among the points. The Blau distances between an individual and his family members, friends, colleagues, and neighbors would be very small in this space, relative to the Blau distances between that individual and dissimilar individuals. In other words, small Blau distances would indicate greater social similarity, while great Blau distances would indicate lesser social similarity. This chapter examines whether there is evidence that socially similar individuals – or Blau "neighbors" – mutually influence each other's political attitudes.

If social similarity does facilitate the mutual influence of political attitudes, we should expect a relationship between Blau distances and political attitudes.

As in the previous chapter, this chapter has two aims – one substantive, the other methodological. The first aim is to see whether socially similar individuals mutually influence each other's political attitudes, all else being equal. To do this, first I elaborate on how social similarity might give rise to social influence. Then, I discuss how to operationalize Blau space and Blau distance and how to construct a spatial weights matrix that would best represent the social connectivities that similar demographic characteristics create among individuals. Next, I use spatial econometric tools to test for spatial influence and to estimate neighborhood influence, and I compare these results with the estimates from a baseline linear regression model and a linear regression model with dummy variables. Based on the results, I discuss whether and to what extent social similarity abets the mutual influence of political attitudes.

The second aim of this chapter is to illustrate the idea of discretization discussed in chapter 3. In chapter 3, we saw that any set of independent variables can be discretized into a set of mutually exclusive, exhaustive, and discrete dummy variables, or disjoint neighborhoods. The discretized linear regression model based on these disjoint neighborhoods is essentially a dummy variable regression model (through the origin). In this chapter, I discretize a set of demographic variables to build a discretized linear regression model. Next, I estimate both types of models using the same data and compare the results. This illustration will clarify further the differences between the discretized linear regression model and spatial regression model for studying social influence.

SOCIAL SIMILARITY AND POLITICAL ATTITUDES

Like geography, demography can facilitate social influence. According to Huckfeldt's (2009) theory of density dependence, social contexts constrain individuals' options for interaction and association with other people. While he focused on such constraints within geographical contexts, as discussed in the previous chapter, Huckfeldt was open to the idea that these constraints can be found in non-geographical contexts[17] as well: "Density dependence [is] defined with respect to the compositional and distributional properties of the social contexts within which actors are imbedded, where the social context is conceived relative to either a *geographically or non-geographically based population* [emphasis added]." As an example of non-geographical context, he explained that individuals who spend their time in environments full of liberals are less likely to interact with conservatives, and hence they are less likely to be persuaded to adopt conservative views. Huckfeldt's example can easily extend to demographic similarity as well: Individuals who spend their time in environments full of blue-collar workers are less likely to interact with white-collar professionals, and hence they are less likely to be persuaded to adopt views associated with blue collar workers.

The idea that demographically-based constraints affect individual attitudes and behaviors is based on two key assumptions: First, individuals tend to spend more time interacting with socially similar others and less time with socially dissimilar others. Second, demographically-defined social groups (*e.g.,* working class, racial groups) tend to favor certain political policies over others because of (perceived) collective benefits or because of similar value systems (see Lazer, 2010). Social science research suggests that both assumptions are quite reasonable.

[17] Previous studies have considered a variety of non-geographical social contexts, such as contexts based on language similarity (Dow et al., 1984), trade or group membership (Simmons and Elkins, 2004), occupation and township (Lin et al., 2006), trade volume (Beck et al., 2006), and dyadic membership (Beck et al., 2006).

First, individuals do spend more time interacting with those who are socially similar and less time with those who are socially dissimilar. Marsden and Hurlbert (1987) found that discussion partners are more likely to be similar in age, race, and religion. Personal networks tend to be homogenous when it comes to race/ethnicity, age, religion, education, occupation, and gender (McPherson, et al., 2001). Rodden (2009) found that voters tend to be clustered into neighborhoods of similar demographic, occupational, income, and political characteristics. According to Lazer (2010), people tend to have discussions with others who tend to be similar in age, race, religion, and political preferences. These findings are consistent with the well-documented principle of homophily, which refers to the phenomenon that "like attracts like." Because of homophily, there are more social ties among similar individuals and fewer social ties among dissimilar individuals.

Second, the relationship between specific social groups and specific political attitudes has been well-documented. In a study of whether individuals are influenced by the characteristics of the surrounding population, Langton and Rapoport (1975) found that working class residents of Santiago were more likely to support Salvador Allende if they lived in working class areas of the city then in other areas of the city. Huckfeldt and Kohfeld (1989) observed a correlation between race and party support among Mississippi voters in the 1984 presidential election. According to Cho (2003), political candidates often appeal to specific racial or ethnic groups. This suggests that certain political attitudes and behaviors are more closely associated with some racial or ethnic groups than others.

If there is greater association among socially similar individuals and if social groups tend to have certain political attitudes and behaviors, then we should not expect that political attitudes to be randomly dispersed across individuals – instead, they should cluster among individuals who are socially similar.

Why would associating with those who are socially similar others lead to social influence? Impersonal and interpersonal interactions with those who are socially similar can affect the marginal utility that an individual derives from holding a certain political attitude. As a result of these interactions, an individual's political attitude can change. These interactions can be classified into two types of mechanisms of social influence: *coercion* and *learning*.

In demographic contexts, *coercion* (which can be hard or soft) includes social pressure and political mobilization. While individuals might experience *social pressure* to conform to certain attitudes or behaviors or they might be *politically mobilized* to hold certain attitudes or behaviors within geographical contexts, there is reason to believe that social pressure and political mobilization can extend beyond mere geography. In an age of instantaneous communication and high mobility norms, individuals may hang onto their social identities as least as strongly as their geographical identities. For one thing, there may be a stronger psychological attachment to their social identities due to familial and social ties. For example, an individual might strongly identify as a Catholic in part because his family and friends are also Catholic. Also, social similarity may breed more intimate relationships compared with the casual day-to-day interactions that take place in geographical contexts. Ties based on social similarity are more likely to be stronger than ties formed by mere geographical proximity. As a result, the pressure to conform to certain political attitudes and behaviors may be stronger coming from those who are socially similar.

As for political mobilization, political candidates often explicitly target certain electoral groups based on their race or ethnicity. While these candidates may travel to geographical areas with strong concentrations of a certain social group, their appeals are not limited to those areas. They may appeal to all Latino voters or all working class voters in their attempts to show why or how their work has provided collective benefits to these groups. Thus, we may expect political attitudes to cluster around some of these social identities.

Learning is a class of social influence mechanisms that involve information exchange and processing. Learning includes *persuasion, socialization*, and *information processing*. First, individuals are more open to persuasion when their persuaders are socially similar to them. According to Lupia (2002), an important element of persuasion is the personal attributes of those attempting to persuade. Those attempting to persuade need to be (or appear to be) trustworthy and knowledgeable to even call attention to an argument. When the argument is heard, whether individuals are receptive to the argument depends (to some extent) on the personal ethos of those presenting the argument. One would expect that arguments tend to be more persuasive if they come from those who are socially similar because they tend to share the same background, values, and/or history, which tends to breed trust.

Second, socialization depends on the existence of social ties to transmit values and attitudes. Social science research has shown that this transmission often takes place through conversation. Huckfeldt and Sprague (1995) found evidence that an individual's voting preferences were closely related to the voting preferences of those with whom he discussed politics, even after controlling for a variety of individual-level variables. This led them to conclude that "the political preferences of citizens have important consequences for the vote choices of other citizens who look to them as political discussants." In a study of the 1992 British General Election, Pattie and Johnston (2000) found that political conversations were a key mechanism in transmitting social influence and concluded:

> Conversations with a party's supporters encouraged respondents to vote for it too, and discouraged them from voting for other parties, especially if those conversations

> took place within their families. … [F]amilies who talked together (more or less) voted together.

Since social ties are often based on social similarity (*i.e.*, the principle of homophily), the values and attitudes transmitted to individuals not are a random sample of all the possible values, attitudes, and information. Accordingly, we should expect individuals to be exposed only a subset of all possible values and attitudes, and that this subset is determined, in part, by the demographic characteristics of the individuals and their social ties.

Finally, the transmission of values and attitudes through social ties is not limited to the transmission of political agreement. This is because social ties are also important for transmitting political information. According to McClurg (2006), "politically sophisticated" social networks can also provide political expertise, increase exposure to norms of political involvement, and communicate political information with more clarity and context. But these functions need not be limited to politically sophisticated social networks. Socially similar individuals may not necessarily provide political expertise, but they might provide information such as the relevance of elections for different social groups, how the policies of different political candidates and parties may benefit or harm one social group over another, and what issues are most salient to which groups. As Huckfeldt (1983b) points out, "[T]he consequences of the interaction need not be explicitly political in order to have political implications." Individuals can use group identification and group norms to process political information and as a result, individuals who are socially similar may come to have similar political points of view.

The mechanisms of coercion and learning are not contradictory or mutually exclusive; indeed, they may work side by and side and reinforce each other. Furthermore, because individuals are themselves members of their demographic contexts, the social influence arising from coercion and learning should be considered a dynamic process. An individual is affected by other socially similar individuals, but those other individuals are also affected by what that individual does as well. Once an individual adopts an attitude or engages in a behavior, the process of social influence does not stop there; that individual also affects others in the same demographic context. Individuals may be pressured socially into adopting a certain political attitude, but once they do so, they themselves become part of that social pressure. Individuals who are politically mobilized to adopt a certain political attitude strengthen the association between the political attitude and social group. After acquiring political information filtered through their social ties, individuals might pass the same filtered information onto others and affect their political attitudes.

Recognizing the relevance of demographic characteristics, political scientists usually include them as control variables in their regression models. In contrast, this chapter looks at demographic characteristics in a more meaningful way. By treating demographic characteristics as the dimensions of a Blau space, one can measure the social proximity among individuals and see if socially-proximate individuals do mutually influence each other's political attitudes. In this way, this chapter puts homophily to task as the organizing principle of social influence.

At the same time, we cannot suppose that demography is destiny. Political attitudes are meaningful so long as they are tied to actual values and ideas, which are ultimately rooted in individual rumination in light of prevailing attitudes and behaviors. As social identity brushes up against individual identity, social influence is limited by individual thought. While the limits on social influence may be very different from individual to individual, the process of social influence does not continually alter an individual's political attitude, but reaches some kind of equilibrium between the individual and the group. As we will see, the models of social influence in this chapter incorporate this process of mutual influence and equilibrium.

METHODOLOGICAL APPROACH

To see whether socially similar individuals mutually influence each other's political attitudes, I first operationalize the idea of social similarity by defining the dimensions of a Blau space. Next, I measure the Blau distances between each pair of respondent by using what is known as *Gower's distance*, and use these distances to construct spatial weights matrices. Next, I use the spatial weights matrices to test for a relationship between the respondents' Blau locations and their political attitudes. Finally, I test whether that relationship holds after controlling for individual-level explanatory and control variables in spatial lag models of influence. To illustrate the discretization process discussed in chapter 3, I compare the results of the spatial lag model with estimates from corresponding discretized linear regression models.

Dependent and Independent Variables

This chapter focuses on only two political attitudes from the 2004 ANES.[18] They are the respondents' self-placements on the following two seven-point political issue scales:

1. Whether the federal government should let everybody get ahead on their own or ensure that every person has a job and a good standard of living
2. Whether the government should provide fewer services to reduce spending or more services even if it means more spending

[18] While I estimated models for all eight political issues, only two featured statistically significant neighborhood effects. To emphasize the methodological points in this chapter, I therefore limited the analysis in this chapter to these two.

The independent variables are explanatory variables that may be relevant to political attitude formation. The explanatory variables are the following eight individual-level variables:

1. Whether the individual is married[19]
2. Religiosity (whether the respondent attends church)
3. Intelligence (based on the interviewer's assessment on a 5-point scale, where higher values correspond to greater intelligence)
4. Household income
5. Level of education (based on seven possible levels; 0 indicates no high school diploma or less, 7 indicates an advanced degree)
6. Level of interest in political campaigns (based on a 3-point scale, where higher values correspond to greater interest)
7. Level of political knowledge (based on the number of correct responses to three knowledge questions)
8. Party identification (based on a 7-point self-placement scale, where higher values correspond to stronger identification with the Republican Party and lower values correspond to stronger identification with the Democrat Party)
9. Political ideology (based on a 7-point self-placement scale, where higher values correspond to greater conservatism and lower values correspond to greater liberalism)
10. The importance of the political issue (on a 5-point scale, where higher values correspond with greater importance to the respondent)

Dimensions of the Blau Space

Previous research shows that individuals tend to share certain demographic characteristics with their friends and family members. Using network data from the 1985 General Social Survey, Marsden (1988) found that respondents and members of their

[19] Marital status is treated as an indicator of social investment rather than a demographic variable.

discussion networks tend to share similar demographic attributes, such as age and race. In a subsequent study, McPherson et al. (2006) compared network data from the 1985 General Social Survey with network data from the 2004 General Social Survey and found that the heterogeneity of confidantes in terms of age, education, race, and sex has remained relatively stable. Based on these studies, I will use age, education, race, and sex as dimensions of the Blau space for this chapter.

One possible concern about this choice of Blau dimensions is that the characteristics of age, education, race, and sex are too general. In politics, social similarity matters in general, but specific types of social similarity matters more in specific cases. For example, British election studies point to the importance of social class, while U.S. race relations studies point to the importance of race. In this chapter, however, the dependent variables are the political attitudes toward the issue of the role of the government in securing jobs and a good standard of living and the issue of government spending and services, which are quite general. Thus, it is appropriate to use a more general definition of social similarity.

Discretization of Demographic Variables

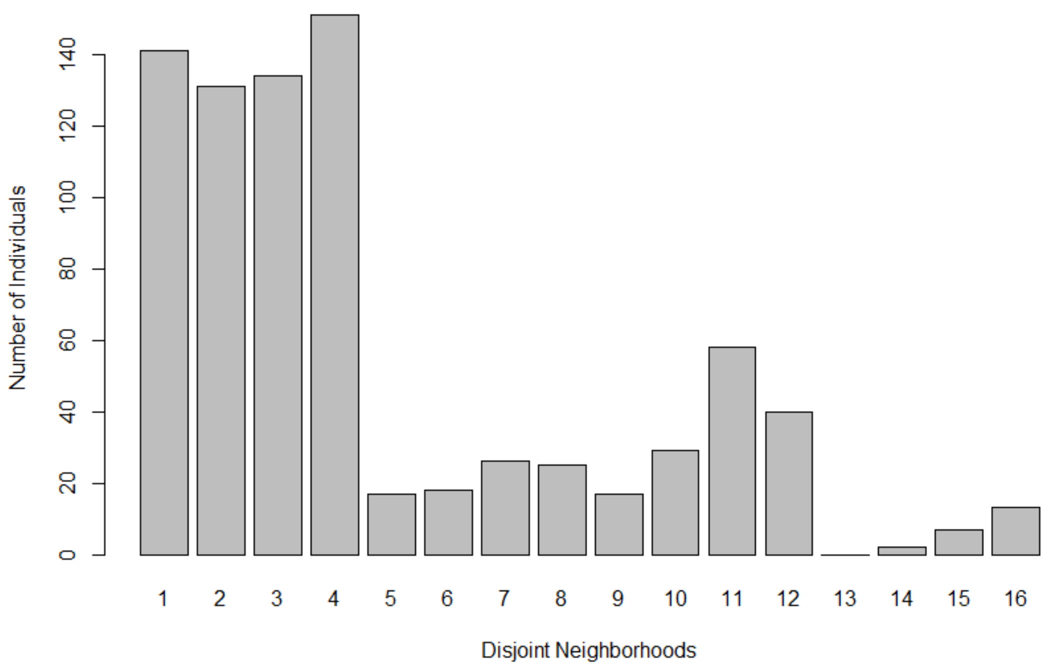
We have seen in chapter 3 that any continuous independent variable can be discretized into p mutually exclusive and exhaustive categories so that a multiple regression model with k continuous variables can be converted into a discretized linear regression model with p^k dummy variables. This section applies this technique to the demographic variables of the Blau space.

The demographic variables that make up the Blau space used in this chapter are age, sex, education, and race. For the sake of simplicity and to keep the number of dummy variables under control, I dichotomize each variable so that $p = 2$ for all $k = 4$ variables in the following manner:

- *Age*: Under 30 (young adult) or 30 and over (adult)
- *Sex*: Female or male (unchanged)
- *Education*: College degree or no college degree
- *Race*: Black or other race

Dicotomizing the five demographic variables yields 16 dummy variables, which represent 16 disjoint neighborhoods. Each disjoint neighborhood represents a unique combination of demographic traits, and each individual is considered a member of one neighborhood only and no other. Using data from the 2004 ANES, a histogram of the 16 disjoint neighborhoods is given in Figure 6.1 below. The number of individuals in each disjoint neighborhood ranges from 1 to 151. The largest neighborhood is neighborhood 4, whose 151 members are white females who are aged 30 and over and have no college degree. The smallest neighborhood is neighborhood 14, which has 2 individuals who are black females who are under 30 years old and have college degrees. Neighborhood 13 is defined by individuals who are black, male, under 30 years old, and have college degrees, but it has no residents.

Figure 6.1. Histogram of 32 Disjoint Neighborhoods



** Each number on the x-axis denotes a disjoint neighborhood.*

Measuring Social Similarity

To measure social similarity, I calculated Gower's dissimilarity coefficient for each pair of respondents. Gower's dissimilarity coefficient *S* is a measure of dissimilarity that is widely used in cluster analysis and multidimensional scaling (see Gower, 1971 and Kaufman and Rousseeuw, 1990). Like other measures of dissimilarity, Gower's dissimilarity coefficient expresses the dissimilarity between two individuals in terms of a distance function based on their demographic characteristics. This distance function can be used to group individuals together based on their similarities and dissimilarities. The advantage of using Gower's dissimilarity coefficient is that it is well-suited for dealing with mixed data types. Since the characteristics (*i.e.*, age, sex, education, and race) that comprise the Blau space in this chapter are of different data types, Gower's coefficient is an appropriate choice.

Gower's S_{ij} for two individuals i and j is the weighted mean of the contributions of their N characteristics. In general, the dissimilarity between two individuals i and j on characteristic m ($m = 1, 2, \ldots, N$) is represented by a score s_{ijm}, which is equal to zero when they are identical and one (for nominal or binary variables) or a fraction[20] otherwise (for ordinal variables). The weight δ_{ijm} is 1 when the comparison is valid and 0 otherwise. Gower's coefficient S_{ij} is given by

$$S_{ij} = \frac{\sum_{N} \delta_{ijm} s_{ijm}}{\sum_{N} \delta_{ijm}}$$

Gower's S_{ij} is bounded between zero and one; higher values indicate greater dissimilarity (or greater social distance), while lower values indicate lesser dissimilarity (or lesser social distance).

Using 2004 ANES data and the cluster package in [R], I computed Gower's S_{ij} for 848 respondents. After accounting for missing values, this resulted in 359,128 dissimilarity scores, with a mean of 0.44 and standard deviation of 0.19.

The Spatial Weights Matrix

Specifying the spatial weight weights matrix is important in spatial econometric analysis. While Gower's distance can provide a measure of the dissimilarity between pairs of respondents, there is no rule to follow, no obvious point at which the similarity between them leads to a social tie. The connectivities between individuals are assumed to be known,

[20] This fraction is the absolute difference of individual *i*'s and *j*'s values for characteristic *m*, divided by the range for that characteristic. See the notes for the `daisy` function in the **`cluster`** package in [R] for more details.

not estimated; the complication is that there are many specification options. Here, I consider two general definitions of social similarity and their corresponding spatial weights matrices. I argue that a spatial weights matrix based on a similarity measure based on Gower's distance is most appropriate for the analysis in this chapter.

One plausible definition of connectivity is social congruence. Two individuals are neighbors if they have the exact set of social characteristics, which means that individuals mutually influence each other only when their race, sex, religion, social class, and age group are exactly identical. Each combination of these characteristics is a disjoint neighborhood and, from the point of view of a respondent, a difference in one of these characteristics makes another individual as equally irrelevant as a difference in all of these characteristics. This idea of neighborhood conflicts with everyday observations: friends tend to be similar, but not identical to one another. Therefore, this definition of social connectivity is inappropriate for a spatial weights matrix.

A more plausible definition of connectivity is based on a sliding scale of influence. This is based on the idea that individuals are more likely to influence each other's political attitudes when they are more socially similar and are less likely to influence each other when they are less socially similar. In other words, all individuals are neighbors of all the other individuals, but individuals are more "neighborly" if they are closer together in Blau space and individuals are less "neighborly" if they are farther apart in Blau space. Under this definition of connectivity, I constructed a spatial weights matrix based on Gower's dissimilarity coefficient. Ranging from zero to one, Gower's distance is a measure of dissimilarity, where greater values indicate greater social distance. By taking the inverse of the dissimilarity between each pair of individuals, I obtain a measure of similarity where greater values indicate greater social similarity. These values are the elements of a spatial weights matrix (with zeroes on the diagonals), for which the spatial lag for each individual can be seen as the weighted average of the attitudes of the other individuals in Blau space, where the weights are inverse Gower's distance. Substantively, this means that the attitudes of those who are socially closer matter more and the attitudes of those who are socially farther away matter less to an individual.

Hereinafter, I refer to this spatial weights matrix as the inverse social distance matrix.

Models and Methods

Using the inverse social distance matrix described above, I calculate Moran's *I* for each dependent variable to test for spatial autocorrelation, or whether there is a relationship between the respondents' Blau locations and their attitudes toward eight political issues. If there is spatial autocorrelation, then one should expect a statistically significant association between political attitudes and neighboring Blau locations. A statistically significant result for a Moran's *I* test would indicate that there is evidence of spatial dependence for the corresponding political attitude.

If there is spatial autocorrelation for a particular political attitude, I then control for several individual-level explanatory and control variables in a spatial lag model to see whether the spatial autocorrelation still holds. Recall that the spatial lag model takes the following form:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon$$
$$\varepsilon \sim N\left(0, \sigma^2\mathbf{I}\right)$$

where β is a *k* by 1 vector of parameters associated with independent variables **X**, which is an *n* by *k* matrix of personal characteristics that may be relevant to an individual's political attitude, such as intelligence, partisanship, interest in politics, and religiosity. By assumption, the error term ε is normally distributed with a constant variance, and **I** is a *n* by *n* identity matrix. $\mathbf{Wy}$ is the spatial lag of $\mathbf{y}$ and is the average attitude of respondent *i*'s neighbors. ρ is the spatial autoregressive parameter; if $\rho = 0$, then there is no spatial dependence, and the spatial lag model reduces to the classical linear regression model. By estimating ρ, we can see whether there is evidence that socially similar individuals mutually influence each other's political attitudes.

To estimate the spatial lag models, I use maximum likelihood (ML) estimation, which can handle the endogeneity of the *y* term. Using ML estimation is preferable to using OLS because the simultaneity bias can result in biased and inconsistent estimates of the parameters.

For the sake of comparison, I also estimate several linear regression models using generalized least squares. These models include a baseline linear regression model with explanatory variables, a linear regression model with explanatory and control variables, and a discretized linear regression model. The discretized linear regression consists of the explanatory variables and 16 dummy variables derived from the demographic characteristics of age, race, education, and sex. By comparing these three linear regression models with the spatial lag model, we can see more clearly the differences between the linear regression and spatial regression frameworks.

HYPOTHESES

In a (global) Moran's *I* test, the null hypothesis is that political attitudes are randomly dispersed in Blau space. If there is no spatial autocorrelation, there would be no association between political attitudes and Blau locations. Because I believe that socially similar individuals mutually influence each other's political attitudes, I expect to find evidence of a positive relationship between values and Blau locations, or positive spatial autocorrelation.

For the spatial lag models of political attitudes, I test whether the spatial autoregressive parameter ρ is statistically significant after accounting for explanatory variables. The null hypothesis is that $\rho = 0$, which corresponds with the lack of mutual influence among socially similar individuals. I expect to find evidence of a neighborhood effect such that $|\rho| < 1$, but remain agnostic as to whether this effect is positive or negative.

TESTING FOR SPATIAL DEPENDENCE: RESULTS

Based on the results of the Moran's *I* tests (see Table 6.1), there is statistically significant spatial autocorrelation for both political attitudes. Substantively, this means that there is a relationship between an individual's attitude toward the issues of government spending and services and the government's role in securing jobs and a good standard of living.

Table 6.1. Results of Moran's I Tests: Inverse Social Distance Weights Matrix

POLITICAL ATTITUDE	Moran's *i*	STANDARD DEVIATE	P-VALUE
Government spending and services	0.056	4.6	< 0.00
Government's role in securing jobs and a good standard of living	0.073	6.2	< 0.00
	*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

In the next section, I investigate whether the spatial dependence detected in the Moran's *I* tests holds after controlling for political knowledge, partisanship, and other individual-level explanatory variables.

ESTIMATING THE SPATIAL LAG AND LINEAR REGRESSION MODELS: RESULTS

In social science research, it is common to use demographic variables as control variables in models of political phenomenon, even though they are seldom statistically significant. By applying this practice to linear regression models of the two political attitudes, we can see that the age, race, education, and sex variables are not statistically significant at the 0.05 level (see Tables 6.2a-6.2b). By comparing the AIC values, we can see that models with the demographic variables actually fit the data slightly worse than the baseline models without them.

Table 6.2a. Linear Regression Models: Government Spending and Services

	BASELINE LINEAR REGRESSION MODEL	LINEAR REGRESSION MODEL WITH DEMOGRAPHICS
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	6.2*** (0.38)	6.1*** (0.4)
Age	-	0.0014 (0.0034)
Female	-	0.11 (0.11)
Race: black	-	0.31* (0.17)
Race: black and Hispanic	-	1.7* (1.0)
Race: black and white	-	1.7* (1.0)
Race: Asian	-	0.24 (0.34)
Race: Asian and white	-	-0.71 (0.83)
Race: Native American	-	0.67 (0.51)
Race: Native American and white	-	0.27 (0.71)
Race: Hispanic	-	0.012 (0.21)
Race: Hispanic and white	-	-0.087 (0.71)
Race: other	-	-0.51 (0.71)
Attends church	0.20* (0.11)	0.16 (0.11)
Intelligence	-0.078 (0.078)	-0.077 (0.079)
Household income	-0.013 (0.0098)	-0.012 (0.010)
Education	-0.029 (0.039)	-0.035 (0.039)
Interest in campaigns	-0.13 (0.085)	-0.13 (0.087)
Political knowledge	-0.13*** (0.045)	-0.11** (0.031)

Table 6.2a. Linear Regression Models: Government Spending and Services

(continued)		
Party ID	-0.14*** (0.029)	-0.11*** (0.031)
Political ideology	-0.35*** (0.046)	-0.37*** (0.048)
Issue importance	0.30*** (0.056)	0.27*** (0.057)
AIC	*2804*	*2815*
Log likelihood	*-1391*	*-1385*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

Table 6.2b. Linear Regression Models: Government's Role in Securing Jobs and a Good Standard of Living

	BASELINE LINEAR REGRESSION MODEL	LINEAR REGRESSION MODEL WITH DEMOGRAPHICS
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	2.2*** (0.042)	2.1 (0.45)
Age	-	0.0068* (0.0037)
Female	-	-0.11 (0.12)
Race: black	-	-0.58*** (0.19)
Race: black and Hispanic	-	-1.5 (0.1)
Race: black and white	-	-0.22 (1.1)
Race: Asian	-	-0.46 (0.39)
Race: Asian and white	-	-0.26 (0.94)
Race: Native American	-	-0.62 (0.57)
Race: Native American and white	-	-1.8** (0.81)
Race: Hispanic	-	-0.21 (0.24)
Race: Hispanic and white	-	-0.31 (0.81)
Race: other	-	-0.21 (0.66)
Attends church	-0.22* (0.12)	-0.19 (0.13)
Intelligence	0.15* (0.086)	0.14* (0.086)
Household income	0.032*** (0.011)	0.027** (0.011)
Education	-0.077* (0.044)	-0.058 (0.044)
Interest in campaigns	0.12 (0.096)	0.088 (0.097)

Table 6.2b. Linear Regression Models: Government's Role in Securing Jobs and a Good Standard of Living

(continued)

Political knowledge	0.099* (0.051)	0.048 (0.053)
Party ID	0.24*** (0.033)	0.20*** (0.035)
Political ideology	0.34*** (0.053)	0.34*** (0.054)
Issue importance	-0.29*** (0.065)	-0.26*** (0.065)
AIC	*3080*	*3081*
Log likelihood	*-1529*	*-1517*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

While the demographic characteristics of age, race, education, and sex were not statistically or substantively significant as control variables in the linear regression models, they do become important as dimensions of a Blau space. Tables 6.3a-6.3b give the results of spatial lag models using the inverse distance matrices based on such a Blau space. These results show that there is a statistically significant neighborhood effect (at the 0.01 level) for the attitude toward the issue of the government's role in securing jobs and a good standard of living ($\hat{\rho} = 0.25$), and a statistically significant neighborhood effect (at the 0.10 level) for the attitude toward the issue of government spending and services ($\hat{\rho} = 0.14$). Substantively, this means that social similarity – conceived and implemented here in terms of similarity in age, race, education, and sex – matters; there is evidence that socially similar individuals mutually influence each other's attitudes toward the issue of government spending and service and the issue of the government's role in securing jobs and a good standard of living.

Table 6.3a. Spatial Lag and Discretized Linear Regression Models: Government Spending and Services

	SPATIAL LAG MODEL	LINEAR REGRESSION MODEL WITH DEMOGRAPHICS
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	5.5*** (0.55)	-
Attends church	0.20* (0.11)	0.17 (0.11)
Intelligence	-0.079 (0.077)	-0.074 (0.078)
Household income	-0.013 (0.0097)	-0.013 (0.010)
Education	-0.020 (0.039)	0.018 (0.074)
Interest in campaigns	-0.13 (0.085)	-0.16* (0.087)
Political knowledge	-0.13*** (0.045)	-0.12** (0.047)
Party ID	-0.14*** (0.029)	-0.12*** (0.031)
Political ideology	-0.35*** (0.046)	-0.35*** (0.048)
Issue importance	0.29*** (0.056)	0.28*** (0.057)
1) ≥ 30 years old, white, college-educated male	-	5.8*** (0.56)
2) ≥ 30, white, college-educated female	-	5.9*** (0.55)
3) ≥ 30 white, non-college educated male	-	6.0*** (0.43)
4) ≥ 30 white, non-college educated female	-	6.2*** (0.43)
5) ≥ 30 years old, black, college-educated male	-	6.8*** (0.63)
6) ≥ 30 years old, black, college-educated female	-	5.6*** (0.64)
7) ≥ 30 years old, black, non-college-educated male	-	6.3*** (0.52)

Table 6.3a. Spatial Lag and Discretized Linear Regression Models: Government Spending and Services

(continued)		
8) ≥ 30 years old, black, non-college-educated female	-	6.1*** (0.51)
9) < 30 years old, white, college-educated male	-	5.9*** (0.62)
10) < 30 years old, white, college-educated female	-	6.2*** (0.59)
11) < 30 years old, white, non-college-educated male	-	5.9*** (0.45)
12) < 30 years old, white, non-college-educated female	-	5.9*** (0.46)
14) < 30 years old, black, college-educated female	-	6.4*** (0.98)
15) < 30 years old, black, non-college-educated male	-	6.5*** (0.70)
16) < 30 years old, black, non-college-educated female	-	6.6*** (0.56)
Spatial autoregressive parameter (ρ)	0.14* (0.085)	-
AIC	*2803*	*2817*
Log likelihood	*-1390*	*-1383*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

Table 6.3b. Linear Regression Models: Government's Role in Securing Jobs and a Good Standard of Living

	SPATIAL LAG MODEL	DISCRETIZED LINEAR REGRESSION MODEL
INDEPENDENT VARIABLES	b (se)	b (se)
Intercept	1.1** (0.52)	-
Attends church	-0.20* (0.12)	-0.19 (0.13)
Intelligence	0.14* (0.084)	0.14 (0.086)
Household income	0.030*** (0.011)	0.029 (0.011)
Education	-0.080* (0.043)	-0.14 (0.082)
Interest in campaigns	0.11 (0.095)	0.13 (0.097)
Political knowledge	0.082 (0.050)	0.063 (0.052)
Party ID	0.23*** (0.032)	0.20 (0.035)
Political ideology	0.34*** (0.052)	0.37 (0.054)
Issue importance	-0.27*** (0.064)	-0.27 (0.066)
Spatial autoregressive parameter (ρ)	0.25*** (0.081)	-
1) ≥ 30 years old, white, college-educated male	-	2.8*** (0.61)
2) ≥ 30, white, college-educated female	-	2.7*** (0.60)
3) ≥ 30 white, non-college educated male	-	2.5*** (0.47)
4) ≥ 30 white, non-college educated female	-	2.4*** (0.47)
5) ≥ 30 years old, black, college-educated male	-	2.6*** (0.69)
6) ≥ 30 years old, black, college-educated female	-	2.1*** (0.71)

Table 6.3b. Linear Regression Models: Government's Role in Securing Jobs and a Good Standard of Living

(continued)		
7) ≥ 30 years old, black, non-college-educated male	-	2.0*** (0.56)
8) ≥ 30 years old, black, non-college-educated female	-	2.2*** (0.55)
9) < 30 years old, white, college-educated male	-	2.6*** (0.69)
10) < 30 years old, white, college-educated female	-	2.3*** (0.64)
11) < 30 years old, white, non-college-educated male	-	2.4*** (0.50)
12) < 30 years old, white, non-college-educated female	-	2.5*** (0.50)
14) < 30 years old, black, college-educated female	-	2.3* (1.3)
15) < 30 years old, black, non-college-educated male	-	0.31 (0.76)
16) < 30 years old, black, non-college-educated female	-	1.5** (0.63)
AIC	*3073*	*3085*
Log likelihood	*-1524*	*-1518*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

Alongside the results for the spatial lag models are the results for the discretized linear regression models. Most of the estimated coefficients for each of the dummy variables are statistically significant at the 0.10 level or better, which means that there is evidence that the 15 categories are important for understanding an individual's baseline political attitude toward the two political issues. For example, individuals who are 30 years old or older, white, college-educated, and male have a mean political attitude of 5.8 toward the issue of government spending and services, which means that – all else being equal – they favor the idea that the government should provide more services rather than less. Through the dummy variables, the discretized linear regression models show that each social category associated with a different mean political attitude for each political issue. But estimating the mean political attitudes of each social group is not the same thing as estimating the influence of belonging to a social group. Using the discretized linear regression model, we cannot say whether an individual who is 30 years old or older, white, college-educated, and male heeds the political attitudes of other individuals who are also 30 years old or older, white, college-educated, and male; we can only say that – assuming that we know nothing else about him – that his political attitude toward the issue of government spending and services is about a 5.8 on a 7-point scale.

Moreover, the discretized linear regression models cannot say whether there is any relationship between the social characteristics of each category and the mean political attitudes. Individuals who are 30 years old or older, white, college-educated, and *male* have a mean political attitude of 5.8 toward the issue of government spending and services, whereas individuals who are 30 years old or older, white, college-educated, and *female* have a mean political attitude of 5.9 toward the issue of government spending and services. This suggests that being male or female makes a difference in political attitudes, but as we can see from the linear regression model with demographic variables (Table 6.2a), sex is not a statistically significant independent variable. By contrasting the discretized linear regression model with the linear regression model with demographic variables, we can observe the loss of information from discretizing independent variables to form disjoint neighborhoods.

Furthermore, by contrasting both linear regression models with the spatial lag model, we can see that the spatial lag model allows demographic characteristics to play a more meaningful role in explaining social influence. In the linear regression models, demographic characteristics are control variables; their associated effects represent little more than adjustments in mean political attitudes. On the other hand, in spatial lag models, demographic characteristics are used to measure the social similarity between each pair of individuals, and these social similarity measurements play a key role in describing how individuals mutually influence each other. Since the spatial weights matrix is specified in terms of inverse social distance, the spatial lag term represents the idea that the attitudes of those who are more socially similar matter more than the attitudes of those who are less socially similar. Since the **y** term is found on both sides of the equation, this social influence process is a process of mutual influence – an individual is more heavily influenced by socially similar others, but he also influences them as well. Thus, the spatial lag models estimated in Tables 6.3a and 6.3b represent a dynamic endogenous process of social influence within Blau space.

The importance of the dynamic endogenous social influence process represented in the spatial lag models can be see more clearly by comparing the total effects of the spatial lag models with the marginal effects of the baseline linear regression models (Tables 6.4a-6.4b). A statistically significant neighborhood effect is a magnifier, which means that the total effect of each independent variable is actually greater than the effect given by the linear regression model. While the marginal effects (1) from the linear regression model are very similar to the direct effects (2) from the spatial lag model, they are quite different from the total effects (4). The total effects consist of direct effects from the individual-level independent variables as well as the indirect effects resulting from the mutual influence process.

Table 6.4a. Comparison of Effects for the Issue of Government Spending and Services

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL		
	(1)	(2)	(3)	(4)
	b	(Direct Effect)	Indirect Effect	Total Effect
Attends church	0.20	0.20	0.032	0.23
Intelligence	-0.079	-0.079	-0.013	-0.092
Household income	-0.013	-0.013	-0.002	-0.015
Education	-0.020	-0.020	-0.0033	-0.024
Interest in campaigns	-0.13	-0.13	-0.021	-0.15
Political knowledge	-0.13	-0.13	-0.020	-0.15
Party ID	-0.14	-0.14	-0.022	-0.16
Political ideology	-0.35	-0.35	-0.056	-0.41
Issue importance	0.29	0.29	0.046	0.33

Table 6.4b. Comparison of Effects for the Issue of the Government 's Role in Securing Jobs and a Good Standard of Living

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL		
	(1)	(2)	(3)	(4)
	b	b (Direct Effect)	Indirect Effect	Total Effect
Attends church	-0.22	-0.20	-0.068	-0.27
Intelligence	0.15	0.14	0.048	0.19
Household income	0.032	0.030	0.010	0.040
Education	-0.077	-0.080	-0.027	-0.11
Interest in campaigns	0.12	0.11	0.038	0.15
Political knowledge	0.099	0.082	0.027	0.11
Party ID	0.24	0.23	0.077	0.31
Political ideology	0.34	0.34	0.11	0.45
Issue importance	-0.29	-0.28	-0.092	-0.37

CONCLUSION

If demographic variables are truly important for understanding political attitudes, it would not suffice to toss in the usual list of such variables in a regression model. Doing so would not capture the most theoretically interesting aspect of these characteristics – that race/ethnicity, age, education, and sex provide social contexts in which individuals think and talk about politics. As McPherson (2004) noted, "The meat grinder of regression puts all the information about social context into the maw of the independent variable, and crunches out values of the dependent variable with the teeth of the estimated parameters, shredding any sense of social structure left in the data." In this chapter, demographic characteristics take on a more meaningful role as the dimensions of a space in which individuals have the opportunity to mutually influence each other.

Using the idea of Blau space, this chapter is based on the view that demographic characteristics define each individual's social position relative to one another. This allows the social similarity between a pair of individuals to be measured as a distance in the Blau space. I used these distances to specify spatial weights matrices to test for spatial autocorrelation in political attitudes toward the issues of government spending and services and the government's role in securing jobs and a good standard of living. These tests showed that there was positive spatial autocorrelation in these two political attitudes. The spatial weights matrices were then used in spatial lag models, where the spatial lag was defined as the weighted average of every other individual's political attitude. Because the weights were based on inverse social distance, the model represents the idea that the attitudes of those who are more socially similar matter more than the attitudes of those who are less socially similar. The statistically significant estimates of the spatial autoregressive parameter ρ for both models are evidence that socially similar individuals do mutually influence each other's political attitudes, all else being equal.

Using demographic variables in this way is a significant departure from the conventional view that demographic characteristics are indicators of group interests or personal attributes that need to be controlled for. In fact, controlling for demographic characteristics can be undesirable; we saw in this chapter that including these characteristics as control variables in a linear regression model of political attitudes actually worsened the fit of the models to the data, demonstrating that a linear regression model with demographic control variables was little more than a "garage can" model.

In addition to showing that socially similar individuals mutually influence each other's political attitudes, this chapter demonstrated a methodological point: that using demographic variables in a linear regression model is not a sufficient substitute for using a spatial lag model of social influence. Following the process of discretization described in chapter 3, I discretized four demographic variables into 16 dummy variables and used these variables to construct two discretized linear regression models. While estimates for these models showed that while the coefficients associated with the dummy variables were statistically significant, they did not produce information beyond the estimated mean political attitudes of each disjoint neighborhood. In contrast, the spatial lag models contained spatial lags instead of the 16 dummy variables, while the other explanatory variables were the same. Constructed from the demographic characteristics of the respondents, the spatial weights matrices represented the social relationships between respondents. By comparing the results of the spatial lag models with the results of several linear regression models, we saw the spatial econometric approach was more substantively meaningful for understanding social influence because the spatial lag models represented a substantive process of mutual influence and because the spatial lag models allowed for the estimate of the size of the influence from socially similar individuals via the spatial autoregressive parameter ρ.

The next chapter is the third and last application of this dissertation's spatial econometric analysis of social influence. It is an analysis of whether individuals mutually influence each other's political attitudes in another type of non-geographical space: ideological space.

CHAPTER 7: SOCIAL INFLUENCE AND IDEOLOGICAL PROXIMITY

If you call yourself a Republican, how would you think about a new political issue? Some Republicans might begin by reasoning from first principles: Does this political issue infringe upon individual liberty? Does it involve new taxes? Or does it involve the expansion of government, and if so, do the benefits outweigh that cost? Other Republicans might look to other Republicans first. What does Jonah Goldberg think about this new issue? What does Karl Rove think? What do the writers of the *Wall Street Journal* editorial page think? Most people, if they are interested in an issue, probably do a little bit of both. If we expect a Republican to reason according to more conservative principles and/or heed the opinions of more conservative commentators, then we should also expect most Republicans to reach similar conclusions about a political issue.

While the outcome of these two thought processes may be the same and while individuals may engage in one, both, or neither of the two, the processes themselves are different; the first is an individual-level process, while the second is social influence. To understand the role of social influence in political attitudes, it is important to separate the two processes and assess their distinct influences on an individual's political attitude. In this chapter, I examine whether ideologically proximate individuals mutually influence each other's political attitudes, all else being equal.

As in the previous two chapters, this chapter has an additional methodological aim. One of the benefits of using the spatial econometric approach is that it requires researchers to make explicit their assumptions regarding the structure of mutual influence. This is done through the specification of the spatial weights matrix, which represents the connectivities among individuals. In this chapter, I compare the results of using two different yet equally plausible spatial weights matrices. Each matrix is based on a different assumption regarding the structure of mutual influence among individuals in an ideological space. By estimating spatial lag models with different spatial weights matrices, we can see whether the evidence of mutual influence among ideologically similar individuals relies on a narrower or broader set of assumptions regarding the structure of mutual influence.

POLITICAL BEHAVIOR AND IDEOLOGICAL PROXIMITY

Previous studies have found – to no one's great surprise - that ideologically similar individuals find each other more agreeable than do ideologically dissimilar individuals, and consequently, individuals are more likely to discuss politics with those who share their political orientations. According to the two-step flow explanation (see Lazarsfeld et al, 1968 and Huckfeldt and Sprague, 1991), individuals are more likely to obtain political information from other individuals rather than directly from politicians and reporters. Researchers have found that these other individuals do not represent a random sample of information and viewpoints; instead, information seekers tend to seek out individuals who are ideologically similar. Unsurprisingly, Democrats prefer discussing politics with other Democrats while Republicans prefer discussing politics with other Republicans (Huckfeldt and Mendez, 2008).

Since ideologically similar individuals are more likely to discuss politics with one another, ideologically similar individuals are more likely to influence each other than ideologically dissimilar individuals. Studies have shown that greater levels of ideological agreement in discussion networks have greater effects on individual attitudes and behaviors. Huckfeldt and Sprague (1991) found that respondents who reported substantial disagreement with their discussion partners are unaffected by their politics, while respondents who reported a lack of disagreement with their discussion partners were associated with large effects on vote choice. In a study of changes in party identification, Kenny (1994) found evidence of social influence when respondents correctly perceive the political attitude of their discussants and when respondents and discussants agree about politics. In contrast, individuals are less likely to engage in certain political behaviors when their social network members are less ideologically similar. Mutz's (2002) study of cross-cutting networks found that individuals with politically-heterogeneous social networks tended to retreat from political activity to avoid social conflict.

How does ideological proximity facilitate social influence? As theorized in the previous two chapters, interactions with those who are geographically proximate and/or socially similar changes an individual's marginal utility of holding a political attitude. How can ideological proximity also alter this marginal utility?

Ideological proximity can abet social influence chiefly through the mechanism of *learning*. Recall from the previous chapter that learning in demographic contexts involves the transmission of values, attitudes, and information through demographically-based social ties and that this process may not be strictly political to have political consequences. In contrast, learning in ideological contexts is intentionally political and requires a greater degree of initiative from individuals. Accordingly, we should expect that learning arises primarily from *persuasion* and *information processing*.

When it comes to politics, individuals are more open to persuasion when their persuaders are ideologically similar to them. According to Lupia (2002), whether an attempt at persuasion is successful depends in part on whether the would-be persuader appears or is actually trustworthy and knowledgeable. Ideologically similar individuals might find each other more trustworthy and knowledge in the narrow sense that they hold common values and recognize similar tradeoffs in politics. Because they share the same ideological background, arguments in favor of one policy or another need not be bogged down in disagreements over fundamental values; instead, they can address whether a policy supports or contradicts a common set of fundamental values.

By interacting with those who are ideologically similar, individuals can process political information more efficiently. It is well-known in political science that that information acquisition is relatively costly (Downs, 1957) and that political sophistication is generally low among voters (Converse, 1964; Luskin, 1987). Accordingly, when confronted with new political information, an individual may not have all the necessary background information with which to assess it. Accordingly, the politically-motivated but cost-conscious individual has at least two options: First, the individual can seek out those with political expertise; second, the individual can look to see how ideologically similar individuals have processed the same information.

First, individuals may seek out political expertise as a way to cut information costs. By referring to political experts in their networks, individuals have a shortcut to an important source of information to circumvent the costs of information acquisition (McClurg, 2003; McClurg 2006b). This behavior is rational because those possessing a high level of political sophistication are more likely to have more and better political information (Delli Carpini and Keeter, 1996; Huckfeldt, 2001) than those with lower levels of political sophistication. According to McClurg (2006), political experts in an individual's network are better able to provide "*clearer* and more *contextualized* communication of political information":

> Expert discussants are useful to their peers because they add clarity to information exchanges in networks, thereby helping people connect that information to their predispositions. Therefore, people who talk politics with political experts are in a better position to identify, reject, and understand the relevance of dissonant political information exchanged in their networks. The primary consequence of this process should be to reduce ambivalence about and increase confidence in their political views.

Because individuals find ideological similarity more agreeable, however, it is unlikely that the political expertise they look for is ideologically neutral. If individuals seek out political expertise that is ideologically compatible with their views, then the political information they acquire will tend to be contextualized and filtered with regard with their common ideology. Consequently, compared with ideologically dissimilar experts, ideologically similar experts are better able to influence individuals to adopt one political attitude over another.

Second, the attitudes of ideologically similar individuals are themselves important sources of information. Assuming that an individual knows the ideological inclinations of others, he can infer whether an attitude toward a specific political issue is closer or farther away from his ideal point based on whether his ideological peers have adopted that attitude. According to Grossback et al. (2004), adoption by ideologically similar peers is an important informational signal that a policy is a good "ideological fit." This is especially important if a policy is not easily identified with ideological labels such as "liberal" or "conservative." If an individual sees that most of his ideological peers prefer one policy over another, this information can help him gauge whether a policy roughly corresponds with his ideology, even if he does not have a clear or complete idea of what the policy is supposed to accomplish. Thus, the attitudes of ideologically similar peers provide important ideological cues regarding specific political issues, and these ideological cues can help individuals overcome their uncertainty regarding specific policies.

Thus, by aiding persuasion and information processing, ideological similarity can facilitate the mutual influence of individual political attitudes. The idea of mutual influence among ideologically similar individuals supports the idea that politics is primarily a social activity. While it is certainly possible that a given individual may arrive at a particular political attitude unaided by and unexposed to the views of others, individual limitations suggests that it is unlikely for this to be true of all political attitudes.

METHODOLOGICAL APPROACH

The main methodological challenge in this chapter lies in measuring the ideological distances between pairs of individuals. These distances are important for representing the connectivities between individuals in spatial weights matrices for spatial econometric analysis. In this section, I first review the variables used in this chapter's analyses. Next, I discuss the issue of measuring ideological similarity and constructing spatial weights matrices. Finally, I review the models, methods, and hypotheses.

Dependent and Independent Variables

The focus of this chapter is the analysis of eight individual-level political attitudes from the 2004 ANES. These dependent variables are respondents' self-placements on seven-point political issue scales, which are as follows:

1. Whether medical expenses should be covered by government insurance or by private insurance
2. Whether the environment should be protected even if it eliminates jobs or reduces the standard of living
3. Whether the federal government should let everybody get ahead on their own or ensure that every person has a job and a good standard of living
4. Whether the government should provide fewer services to reduce spending or more services even if it means more spending

5. Whether the federal government should make every effort to improve the social and economic position of blacks
6. Whether women should have an equal role with men in business, industry, and government or remain in the home
7. Whether defense spending should be increased or decreased
8. Whether abortions should be more or less restricted

Detailed descriptions and summary statistics for these eight dependent variables can be found in the appendix.

Independent variables include control variables and explanatory variables that may be relevant to political attitude formation. These include the following individual-level socio-demographic (control) variables:

1. Age (in years)
2. Number of children
3. Whether the respondent is female
4. Whether the respondent lives in a rural environment
5. Whether the respondent is married
6. Whether the respondent belongs to the military
7. Social class (self-placed scale, where higher values correspond to higher social classes)

Explanatory variables include the following individual-level variables:

1. Religiosity (whether the respondent attends church)
2. Intelligence (based on the interviewer's assessment on a 5-point scale, where higher values correspond to greater intelligence)
3. Household income
4. Level of education

5. Level of interest in political campaigns (based on a 3-point scale, where higher values correspond to greater interest)

6. Level of political knowledge (based on the number of correct responses to three knowledge questions)

7. Party identification (based on a 7-point self-placement scale, where higher values correspond to stronger identification with the Republican Party and lower values correspond to stronger identification with the Democrat Party)

8. Political ideology (based on a 7-point self-placement scale, where higher values correspond to greater conservatism and lower values correspond to greater liberalism)

9. The importance of the political issue (on a 5-point scale, where higher values correspond with greater importance to the respondent)

Measurement

Before one can assess whether ideologically similar individuals mutually influence each other's political attitudes, it is necessary to define and measure "ideological similarity." At first glance, the political ideology indicator from the 2004 ANES is the most obvious candidate for measuring the ideological similarity between pairs of individuals; we can simply define ideological neighbors as individuals who have similar ideological viewpoints. Unfortunately, measuring ideological similarity may not be as straightforward as one would like. This is because the 2004 ANES and other mass surveys often measure ideology by asking respondents to place themselves on a seven-point scale, where extreme conservatism and extreme liberalism are the endpoints. The problem is that such a scale is too blunt an instrument for measuring the ideological views of respondents. It is clear among political scientists what the left-right spectrum means in politics; it is less clear that laymen know what this spectrum means in relation to their particular set of beliefs (Converse, 1964). Ideally, one would place each individual on some common ideological space free from the limitations of the common seven-point scale.

Fortunately, adherents of the spatial theory of voting have successfully undertaken the task of identifying a common ideological space – or basic space. According to the spatial theory of voting, voters and candidates can be located in an abstract spatial map, and each voter votes for the "closest" candidate on that map. The idea that non-geographical, ideological proximity could motivate political behavior grew out of Hotelling's observation that geographical proximity can motivate economic behavior. Hotelling (1929) showed that in a town where all the houses faced a single road, the best location for a single grocery store was the middle (or median) of the town, and if there were two grocery stores, they would be located right next to each other in the middle of town. Downs (1957) applied this idea of spatial competition to politics in his influential work *An Economic Theory of Democracy*, in which he argued that voters and political parties can be located in common space and that voters vote for the party closest to them, causing the parties to converge to the median.

By leveraging the idea and construction of an ideological basic space in this chapter, I use survey data to map all survey respondents in a common ideological space, calculate their relative distances from each other, and identify ideological neighbors based on these distances. This will allow for the use of spatial econometrics to study the effects of ideological proximity on political attitudes.

The ideological "basic space" refers to a space that comprises a set of latent dimensions common to all voters and politicians. Poole (1998) gives the following definition:

> In standard spatial theory, each issue is modeled as an ordered dimension of alternatives, and each respondent is assumed to have an ideal point on, and single-peaked preferences over, each issue dimension. If the respondents have highly structured belief systems (Converse 1964), then this means that the issues lie on a low-dimensional hyperplane through the issue space.

This low-dimensional issue space has been called a "basic space" (Ordeshook, 1976; Poole, 1998), ideological dimensions (Hinich and Munger, 1994), and "policy space." In this chapter, I will refer to it as basic ideological space.

If we assume that every respondent to the 2004 ANES occupies a position in this space, then we can use Poole's scaling technique to identify those positions based on their responses to certain questions in the survey. Poole's scaling technique for recovering a basic space is based on the assumption that respondents report their ideological locations under different levels of bias and different interpretations of the survey questions. The procedures are based on the singular value decomposition of a numeric matrix with missing data (see Poole, 1998 for mathematical details). The procedures will be briefly described here.

Suppose there are n respondents and m issue scales. The survey data is a matrix X_0, such that each element x_{ij} is individual i's response to the jth issue ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$). Let X denote the version of X_0 with no missing data. Suppose there are s dimensions in the basic ideological space. Let ψ_{ik} represent individual i's position on the kth basic dimension ($k = 1, 2, \ldots, s$). Let W denote an m by s matrix of weights that map the individuals' positions from the basic ideological space to the observed placements on the m issue scales. The model to be estimated is

$$X_0 = [\Psi W' + J_n c']_0 + E_0$$

Since the dependent variables in this dissertation are political issues, it would be tautological to place individuals on a basic space constructed from issue scales. Instead, the transposed version of this model can be used to construct a basic space from respondents' placement of political candidates and political parties on a seven-point liberal-conservative scale. Accordingly, I applied the transposed version of this model to liberal-conservative scales in the 2004 ANES. Survey respondents were asked to place Bush, Kerry, Nader, the

Democrats, and the Republicans on a seven-point liberal-conservative scale. Using Poole's scaling technique and the `basicspace` package in [R], I obtained estimates of the locations of Bush, Kerry, Nader, the Democrats, the Republicans, and, more importantly, the respondents on a two-dimensional basic ideological space. Table 7.1 gives the relevant estimates obtained with the Poole scaling technique.

Table 7.1. Elements of the Basic Ideological Space

STIMULUS	N	COORDINATE 1	COORDINATE 2	r^2
Bush	386	0.565	-0.0.15	0.888
Kerry	386	-0.375	-0.401	0.833
Nader	386	-0.346	0.825	1.000
Democrats	386	-0.373	-0.399	0.823
Republicans	386	0.530	-0.010	0.854

Dimensions estimated: 2

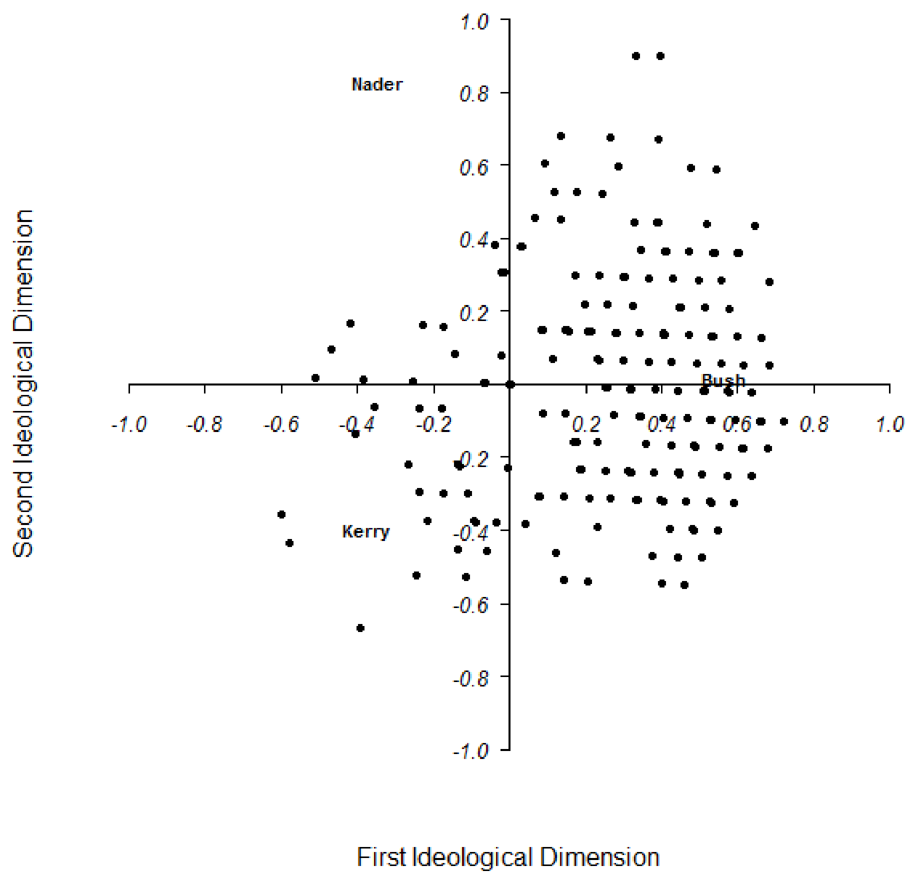
5 rows, 386 columns

Total number of data entries: 1930

Sum of squares: 7497.7

Using these estimates, I map the respondents onto a two-dimensional basic space (Figure 7.1).

Figure 7.1. The Two-Dimensional Basic Ideological Space



Note: There are no labels indicating the ideological points for the Republican Party and Democrat Party because they overlap almost completely with the labels for Kerry and Bush, respectively.

Note that the first ideological dimension (x-axis) is comparable to the conventional 7-point liberal-conservatism scale. The interpretation of the second ideological dimension is less obvious; researchers have speculated that it represents economic policy or support for the status quo.

The Spatial Weights Matrix

In this section, I consider two ways that individuals might be connected to each other in basic ideological space and their corresponding spatial weights matrices. I argue that both are appropriate for representing the connectivities between respondents.

Two plausible definitions of connectivity are inverse ideological distance and k-nearest neighbor. *Inverse ideological distance* is based on the idea that mutual influence is proportional to ideological similarity. This definition of connectivity implies that all individuals are ideological neighbors with one another but that every one's neighborliness is weighted by their ideological similarity in the basic ideological space. In other words, everyone takes account of everyone else's political attitudes, only in different degrees. For a spatial weights matrix based on inverse ideological distance, the spatial lag for each individual can be seen as the weighted average of everyone else in the basic ideological space.

K-nearest neighbor connectivity is based on the idea that an individual is subject to mutual influence only by some ideological neighbors, not all. Like the inverse distance idea, individuals are willing to consider a range of viewpoints, but unlike the inverse distance definition, individuals are willing to consider only a limited range of viewpoints. This is represented by k, which is a constant denoting the number of neighbors for each individual. By convention, the number of neighbors is $k = 5$, which means that an each individual takes (equal) account of the five nearest ideological neighbors and no more. For this spatial weights matrix, the spatial lag for each individual can be interpreted as the unweighted average of five ideological neighbors' attitudes.

Table 7.2 summarizes the two spatial weights matrices. In the following analyses, both types of spatial weights matrices have been normalized so that the row sums equal 1.

Table 7.2. Spatial Weights Matrices for Geographical Proximity

TYPE OF CONNECTIVITY	DEFINITION OF INDIVIDUAL I'S NEIGHBORHOOD	DEFINITION OF SPATIAL WEIGHT W_{IJ}
k-nearest neighbor	The k=5 ideologically closest individuals	$w_{ij} = 1$ if individual j is among individual i's k closest ideological individuals $w_{ij} = 0$ otherwise or $i = j$
Inverse distance	All other individuals	$w_{ij} = d_{ij}$ where d_{ij} = the inverse Euclidean distance between individual i's and j's locations in basic ideological space $w_{ij} = 0$ if $i = j$

Models and Methods

I used the `basicspace` and `spdep` packages in [R] to calculate the distance between respondent *i*'s and respondent *j*'s locations in ideological space and to construct the three spatial weights matrices discussed previously. Using the spatial weights matrices describe in the previous section, I calculated Moran's *I* for each dependent variable to test for spatial autocorrelation. The results of the Moran's *I* tests will show whether there is evidence of spatial dependence. If there is no spatial autocorrelation, there would be no association between values and locations. If there is spatial autocorrelation, then one should expect a statistically significant association between values in a given location with values in neighboring locations. Based on my hypotheses, there should be evidence of a systematic relationship between values and ideological locations, or positive spatial autocorrelation.

Next, I control for individual-level explanatory and control variables in spatial lag models to see whether there is still evidence of spatial dependence. These spatial lag models are based on the inverse distance spatial weights matrix and on the *k*-nearest neighbor spatial weights matrix. The spatial lag model takes the following form:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \varepsilon$$
$$\varepsilon \sim N\left(0, \sigma^2\mathbf{I}\right)$$

β is a *k* by 1 vector of parameters associated with independent variables **X**, which is an *n* by *k* matrix of personal characteristics that may be relevant to an individual's political attitude. The error term ε is assumed to be normally distributed with a constant variance, and **I** is a *n* by *n* identity matrix.

Wy is the spatial lag of **y** and is the average attitude of respondent *i*'s neighbors. Using the inverse distance spatial weights matrix, **Wy** is the weighted average of respondent i's ideological neighbors' attitudes; the attitudes of individuals who are ideologically closer to respondent *i* are weighted more heavily, while individuals who are ideologically farther away are weighted less heavily. Using the *k*-nearest neighbor matrix, **Wy** is the unweighted average of respondent *i*'s $k = 5$ closest ideological neighbors.

The spatial autoregressive parameter ρ indicates the extent of mutual influence among ideologically proximate individuals. If $\rho = 0$, then there is no spatial dependence, and the spatial lag model reduces to the classical linear regression model. If $\rho \neq 0$, then there would be evidence of mutual influence on individuals' political attitudes in ideological space.

To estimate the spatial lag models, I use maximum likelihood (ML) estimation, which can handle the endogeneity of the *y* term in the spatial lag models. Using ML estimation is preferable to using OLS because the simultaneity bias can result in biased and inconsistent estimates of the parameters. I also estimate baseline linear regression models (using generalized least squares) for comparison.

HYPOTHESES

In a (global) Moran's *I* test, the null hypothesis is that political attitudes are dispersed randomly in basic ideological space, and the alternative hypothesis is that the attitudes are not dispersed randomly. Because I believe that individuals mutually influence each other's political attitudes because of their ideological similarity, I expect to find evidence of a positive relationship between individuals' attitudes and locations in basic ideological space, or positive spatial autocorrelation.

For the spatial lag models of political attitudes, I test whether the spatial autoregressive parameter ρ is statistically significant after accounting for explanatory variables. The null hypothesis is that $\rho = 0$, and the alternative hypothesis is that $\rho \neq 0$. I expect to find evidence of a neighborhood effect such that $|\rho| < 1$, but remain agnostic as to whether this effect is positive or negative.

TESTING FOR SPATIAL DEPENDENCE: RESULTS

Based on the results of the Moran's *I* tests (see Table 7.3a-7.3b), both types of spatial weights matrices produced statistically significant (at the 0.10 level or better) spatial autocorrelation for all or most of the political attitudes. The Moran's *I* tests using the inverse distance spatial weights matrices produced statistically significant spatial autocorrelation for seven out of eight political attitudes; the exception was the attitude toward the issue of the role of women. The Moran's *I* tests using the 5-nearest neighbor spatial weights matrices produced statistically significant spatial autocorrelation for all eight political attitudes. Substantively, the Moran's *I* test results mean that there is a positive relationship between an individual's political attitudes and his ideological neighbors' political attitudes.

By comparing the results, we can see that Moran's *I* associated with the 5-nearest neighbor spatial weights matrix is consistently larger than the Moran's *I* associated with the inverse distance matrix. The p-values for the former are also much smaller than the p-values for the latter. Thus, the evidence of spatial autocorrelation is stronger when using the 5-nearest neighbor spatial weights matrix. Because the two types of matrices are based on different assumptions, the stronger autocorrelation for the tests conducted with the 5-nearest neighbor spatial weights matrices suggests that mutual influence may be stronger in smaller pools of ideologically similar individuals.

In the next section, I investigate whether the spatial dependence detected in the Moran's *I* tests holds after controlling for political knowledge, partisanship, and other individual-level independent variables relevant to political attitudes. Because each type of spatial weights matrix produced evidence of spatial autocorrelation, I estimate spatial lag models for all of them and show them alongside a baseline linear regression model for comparison.

Table 7.3a. Results of Moran's I Tests: Inverse Distance Spatial Weights Matrix

POLITICAL ATTITUDE	Moran's *i*	SE	P-VALUE
Access to abortion	0.067***	4.4	< 0.00
Environment versus jobs tradeoff	0.075***	4.9	< 0.00
Government versus private medical insurance	0.043***	2.8	0.0059
Government spending and services	0.047**	3.0	0.0030
Defense spending	0.071***	4.3	< 0.00
Government's role in securing jobs and a good standard of living	0.056***	3.6	0.00037
Government assistance to blacks	0.056***	3.4	0.00057
Role of women	0.020	1.4	0.17
	*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

Table 7.3b. Results of Moran's I Tests: 5-Nearest Neighbor Spatial Weights Matrix

POLITICAL ATTITUDE	Moran's *i*	STANDARD DEVIATE	P-VALUE
Access to abortion	0.15***	6.3	< 0.00
Environment versus jobs tradeoff	0.18***	7.3	< 0.00
Government versus private medical insurance	0.13***	5.5	< 0.00
Government spending and services	0.10***	4.3	0.000017
Defense spending	0.17***	6.8	< 0.0
Government's role in securing jobs and a good standard of living	0.14***	5.6	< 0.00
Government assistance to blacks	0.15***	6.2	< 0.00
Role of women	0.12***	5.1	< 0.00
	*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests		

ESTIMATING THE SPATIAL LAG MODEL: RESULTS

I used maximum likelihood to estimate spatial lag models using the two types of similarity-based spatial weights matrix for eight political attitudes issues. While models for all the dependent variables were estimated, only the results for the political attitudes with statistically significant neighborhood effects are shown (Tables 7.4a-7.4c), alongside estimates for linear regression models for comparison.

The results show that there is statistically significant evidence of spatial dependence for three political attitudes: the attitudes toward the issues of defense spending, environmental protection versus jobs, and access to abortion. The null hypothesis for each spatial lag model of each dependent variable is that there is no neighborhood effect ($\rho = 0$). Based on the results, I reject the null hypothesis for the attitudes toward the issues of defense spending, environmental protection versus jobs, and access to abortion, and I do not reject the null hypothesis for attitudes toward the remaining five issues.

For five political attitudes, the neighborhood effects detected through Moran's *I* tests disappeared after controlling for individual-level characteristics, such as ideology and party identification. These findings suggest that for five political issues, ideologically-similar individuals have similar attitudes because of their similar individual characteristics, not because of mutual influence. In other words, there is no evidence that there is mutual influence in political attitudes within ideologically-defined neighborhoods for these issues.

Based on the results in Tables 7.4a-7.4c, the three political attitudes that are subject to neighborhood effects are the attitudes toward the issues of environmental protection versus jobs, defense spending, and access to abortion. For models of these three attitudes, a comparison of AIC values shows that it is not obvious which spatial lag model fits better. The estimated spatial autoregressive parameter ρ and its standard error are slightly larger for the spatial lag models with inverse distance spatial weights matrix than the corresponding estimates for the spatial lag models with the 5-nearest neighbor spatial weights matrix.

Table 7.4a. Attitude Models: Defense Spending

	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	1.8*** (0.43)	1.2** (0.56)	1.3*** (0.47)
Age	0.0065 (0.0042)	0.0064 (0.0041)	0.0066 (0.0041)
Children	0.012 (0.069)	0.013 (0.068)	0.016 (0.067)
Female	-0.15 (0.12)	-0.16 (0.12)	-0.16 (0.12)
Rural	0.14 (0.14)	0.15 (0.14)	0.15 (0.14)
Married	0.12 (0.11)	0.11 (0.10)	0.10 (0.10)
Military	0.24 (0.16)	0.25 (0.15)	0.23 (0.15)
Attends church	-0.13 (0.12)	-0.13 (0.11)	-0.13 (0.11)
Social class	0.027 (0.036)	0.028 (0.035)	0.031 (0.035)
Intelligence	0.023 (0.086)	0.015 (0.085)	0.0091 (0.085)
Household income	0.015 (0.011)	0.015 (0.011)	0.015 (0.011)
Education	-0.17*** (0.043)	-0.16*** (0.042)	-0.16*** (0.041)
Interest in campaigns	0.016 (0.10)	0.022 (0.10)	0.028 (0.099)
Political knowledge	-0.026 (0.049)	-0.029 (0.048)	-0.034 (0.047)
Party ID	0.20*** (0.033)	0.20*** (0.032)	0.20*** (0.032)

Table 7.4a. Attitude Models: Defense Spending

(continued)			
Political ideology	0.16*** (0.051)	0.15*** (0.050)	0.15*** (0.049)
Issue importance	0.32*** (0.065)	0.32*** (0.064)	0.32*** (0.064)
Spatial autoregressive parameter (ρ)	-	0.17* (0.088)	0.13** (0.057)
AIC	*1713*	*1712*	*1710*
Log likelihood	*-839*	*-837*	*-836*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests			

Table 7.4b. Attitude Models: Environmental Protection versus Jobs

	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	3.8*** (0.54)	3.1*** (0.63)	3.4*** (0.57)
Age	0.0062 (0.0049)	0.0064 (0.0047)	0.0069 (0.0047)
Children	-0.011 (0.081)	-0.0059 (0.079)	-0.00099 (0.078)
Female	0.17 (0.14)	0.15 (0.13)	0.15 (0.13)
Rural	-0.091 (0.17)	-0.10 (0.16)	-0.10 (0.16)
Married	-0.098 (0.12)	-0.10 (012)	-0.097 (0.12)
Military	0.18 (0.18)	0.15 (0.18)	0.13 (0.18)
Attends church	0.036 (0.14)	0.020 (0.13)	0.024 (0.13)
Social class	0.039 (0.042)	0.043 (0.041)	0.038 (0.040)
Intelligence	0.050 (0.10)	0.039 (0.099)	0.045 (0.098)
Household income	0.0022 (0.013)	0.0016 (0.013)	0.0014 (0.013)
Education	-0.17*** (0.050)	-0.16*** (0.048)	-0.16*** (0.048)
Interest in campaigns	-0.031 (0.11)	-0.035 (0.11)	-0.038 (0.11)
Political knowledge	-0.0099 (0.057)	-0.010 (0.056)	-0.0090 (0.055)
Party ID	0.13*** (0.039)	0.13*** (0.038)	0.13*** (0.037)
Political ideology	0.19*** (0.059)	0.17*** (0.058)	0.17*** (0.057)
Issue importance	-0.39*** (0.075)	-0.38*** (0.073)	-0.37*** (0.073)
Spatial autoregressive parameter (ρ)	-	0.25** (0.098)	0.14** (0.063)

Table 7.4b. Attitude Models: Environmental Protection versus Jobs

(continued)

AIC	*1845*	*1842*	*1843*
Log likelihood		*-902*	*-902*

*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests

Table 7.4c. Attitude Models: Access to Abortion

	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
INDEPENDENT VARIABLES	b (se)	b (se)	b (se)
Intercept	3.5 (0.30)	2.9 (0.39)	3.1 (0.34)
Age	-0.0010 (0.0030)	-0.00073 (0.0029)	-0.00065 (0.0029)
Children	-0.11 (0.047)	-0.10 (0.046)	-0.10 (0.046)
Female	0.30 (0.086)	0.30 (0.084)	0.30 (0.084)
Rural	-0.15 (0.10)	-0.14 (0.10)	-0.15 (0.10)
Married	-0.081 (0.074)	-0.089 (0.073)	-0.083 (0.072)
Military	-0.036 (0.11)	-0.019 (0.11)	-0.034 (0.11)
Attends church	-0.40 (0.084)	-0.40 (0.082)	-0.40 (0.082)
Social class	0.013 (0.026)	0.0076 (0.025)	0.010 (0.025)
Intelligence	0.11 (0.062)	0.10 (0.061)	0.10 (0.060)
Household income	0.030 (0.0078)	0.031 (0.0076)	0.031 (0.0076)
Education	0.046 (0.030)	0.045 (0.030)	0.043 (0.029)
Interest in campaigns	-0.0040 (0.035)	-0.014 (0.068)	-0.0099 (0.067)
(continued)			
Political knowledge	0.074	0.073 (0.034)	0.073 (0.034)
Party ID	-0.093 (0.023)	-0.092 (0.022)	-0.093 (0.022)
Political ideology	-0.14 (0.036)	-0.13 (0.035)	-0.13 (0.035)
Issue importance	-0.15 (0.040)	-0.15 (0.039)	-0.15 (0.039)

Table 7.4c. Attitude Models: Access to Abortion

(continued)			
Spatial autoregressive parameter (ρ)	-	0.21** (0.089)	0.12* (0.067)
AIC	*1512*	*1508*	*1510*
Log likelihood	*-738*	*-735*	*-736*
*** p<0.01. ** p<0.05, * p<0.10, 2-tailed tests			

Direct Effects, Indirect Effects, and Total Effects

In the spatial lag model, the presence of the dependent variable y on either side of the equation means that the fitted coefficients must be interpreted in light of direct effects, indirect effects, and total effects, such that the expected value of y is

$$E(y) = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta$$

when there is no neighborhood effect. If there is no neighborhood effect ($i.e., \rho = 0$) then the expected value of y reduces to $\mathbf{X}\beta$.

Tables 7.5a-7.5c contrast the effects found by using the linear regression model with the total effects found by using the spatial lag models. The presence of a statistically significant neighborhood effect means that the total effect of each independent variable is tends to be greater than the effect given by the linear regression model. In Tables 7.5a-7.5c, note that the marginal effects (1) from the linear regression model are very different from the total effects for the spatial lag model with the inverse distance spatial weights matrix (2) and from the total effects for the spatial lag model with the 5-nearest neighbor spatial weights matrix (3). Because the total effect of an independent variable for a spatial lag model is the sum of the direct and indirect effects, the effects indicated in (2) and (3) are generally larger than the effects indicated in (1).

An interesting exception can be found in the comparison of the effects of intelligence on attitudes toward defense spending in Table 7.5a, where the total effects from the spatial lag models are much smaller than the effects garnered from the baselines linear regression model. By referring to the direct effects of intelligence from the spatial lag models in Table 7.4a, we can see that the estimated coefficients (*i.e.*, the direct effects) for intelligence in both spatial lag models are much smaller than the estimated coefficient for intelligence in the baseline linear regression model so that after adding the indirect effects, the total effects from the spatial lag models are still smaller than the effects from the linear regression model. Since intelligence is not a statistically significant variable, I then estimated all three models excluding intelligence as an explanatory variable. The estimated coefficients for all three revised models were essentially the same as the results reported in Table 7.4a; however, the standard errors were slightly smaller and the AIC measures for the spatial lag models were slightly lower. This finding suggests that including intelligence as an independent variable led to a specification error of including an unnecessary variable, which in turn suggests that an inspection of direct, indirect, and total effects may be also be useful as a diagnostic tool.

Table 7.5a. Comparison of Effects for the Issue of Defense Spending

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
	(1)	(2)	(3)
	Effect	Total Effect	Total Effect
Age	0.0065	0.0077	0.0076
Children	0.012	0.016	0.018
Female	-0.15	-0.19	-0.19
Rural	0.14	0.18	0.17
Married	0.12	0.14	0.12
Military	0.24	0.29	0.27
Attends church	-0.13	-0.15	-0.15
Social class	0.027	0.034	0.036
Intelligence	0.023	0.018	0.010
Household income	0.015	0.019	0.018
Education	-0.17	-0.19	-0.19
Interest in campaigns	0.016	0.026	0.032
Political knowledge	-0.026	-0.035	-0.039
Party ID	0.20	0.24	0.23
Political ideology	0.16	0.17	0.17
Issue importance	0.32	0.39	0.37

Table 7.5b. Comparison of Effects for the Issue of Environmental Protection versus Jobs

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
	(1)	(2)	(3)
	Effect	Total Effect	Total Effect
Age	0.0062	0.0085	0.0080
Children	-0.011	-0.0078	-0.0012
Female	0.17	0.20	0.17
Rural	-0.091	-0.14	-0.12
Married	-0.098	-0.13	-0.11
Military	0.18	0.20	0.16
Attends church	0.036	0.026	0.027
Social class	0.039	0.057	0.045
Intelligence	0.050	0.052	0.053
Household income	0.0022	0.0021	0.0016
Education	-0.17	-0.22	-0.19
Interest in campaigns	-0.031	-0.047	-0.045
Political knowledge	-0.0099	-0.014	-0.010
Party ID	0.13	0.17	0.15
Political ideology	0.19	0.23	0.20
Issue importance	-0.39	-0.50	-0.43

Table 7.5c. Comparison of Effects for the Issue of Access to Abortion

INDEPENDENT VARIABLES	BASELINE LINEAR REGRESSION MODEL	SPATIAL LAG MODEL (INVERSE DISTANCE)	SPATIAL LAG MODEL (5-NEAREST NEIGHBOR)
	(1)	(2)	(3)
	Effect	Total Effect	Total Effect
Age	-0.0010	-0.00092	-0.00074
Children	-0.11	-0.13	-0.12
Female	0.30	0.38	0.34
Rural	-0.15	-0.18	-0.17
Married	-0.081	-0.11	-0.094
Military	-0.036	-0.025	-0.039
Attends church	-0.40	-0.50	-0.45
Social class	0.013	0.0096	0.012
Intelligence	0.11	0.13	0.12
Household income	0.030	0.039	0.035
Education	0.046	0.057	0.049
Interest in campaigns	-0.0040	-0.017	-0.011
Political knowledge	0.074	0.093	0.083
Party ID	-0.093	-0.12	-0.11
Political ideology	-0.14	-0.17	-0.15
Issue importance	-0.15	-0.20	-0.17

CONCLUSION

Is there is evidence that ideologically-similar individuals mutually influence each other's political attitudes? Based on this chapter's analysis, the answer is yes. Furthermore, this answer does not depend on a particular assumption regarding the relationships among individuals in ideological space.

In this chapter, I looked at ideology as a space in which individuals might mutually influence each other's political attitudes. The motivating theory is that ideologically similar individuals are important resources for political learning. First, individuals are more receptive to political ideas when they come from ideologically similar individuals, and that this receptivity is social in nature and separate from the ideological leanings of the individuals themselves. Second, individuals tend to circumvent informational costs by seeking out political expertise from ideologically similar experts rather than ideologically neutral ones, leading to a biased filtering and contextualization of political information. Third, the adoption of a particular political attitude by ideologically similar individuals is an important cue regarding the ideological compatibility of a political policy or issue.

In order to assess whether ideologically similar individuals mutually influence each other's political attitudes, I operationalized the idea of "ideological similarity" by measuring the distance between individuals in a common ideological space. To get this space, I used survey respondents' placements of political candidates and political parties to estimate a basic ideological space in which all the respondents are located. After calculating the distances between individuals in the space, I constructed two types of spatial weights matrices based on two definitions of connectivity.

To find out whether political attitudes are contagious, I used Moran's *I* tests to undercover evidence of spatial dependence, and then used spatial lag models to estimate neighborhood effects. The results showed that for three political issues, ideologically similar individuals mutually influence each other's political attitudes. For these three issues, the spatial clustering of political attitudes cannot be explained only by individual characteristics such as party identification and political ideology. Furthermore, the fact that the results were similar for both specifications of the spatial weights matrix shows that the evidence of mutual influence does not depend on a single assumption regarding the structure of influence.

As in geographical and social contexts, there is evidence that of mutual influence in ideological space. Individuals do influence and are influenced by those who are ideologically similar.

CHAPTER 8: CONCLUSIONS

This dissertation is an attempt to explain how spatial econometrics can be used to study social influence in politics. I have explained how social influence is an endogenous process of mutual influence among similar individuals, where "similarity" in terms of social distance can be based on geographic, demographic, or ideological characteristics. I have discussed the limitations of conventional models for studying social influence, such as ANOVA, linear regression, and contextual models, showed how these models relate to spatial econometric models, and discussed the advantages of using spatial regression models over these conventional models. I have argued that spatial econometrics can be a feasible tool for studying social influence by bringing together the concepts of social influence and neighborhood effect and by broadening the concept of space beyond mere geography to a more general set of opportunities for mutual influence based on characteristics such as geographical proximity, demographic similarity, and ideological similarity. Using the 2004 ANES, I applied these ideas by developing spatial regression models for individual attitudes toward political issues to see whether individuals who are geographical, demographic, and ideological similar mutually influence each other's attitudes. I summarize these contributions below.

SUMMARY OF FINDINGS

The theoretical feasibility of using spatial econometrics for studying social influence in politics

In chapter 2, I argued that social contexts can be conceived as multi-dimensional spaces that provide opportunities for social influence. Social influence should be understood as mutual influence among proximate individuals in such a space. This dissertation considered three types of spaces: geographical, demographic (Blau), and ideological space. These spaces provide opportunities for social influence in the sense that individuals are more likely to associate with those who geographically proximate, demographically similar, and/or ideologically similar.

Understanding social context as space offers two main advantages. First, it incorporates previous definitions of social context (*e.g.*, census tract) while providing a flexible framework for more abstract definitions (*e.g.*, socioeconomic status, network membership) that need not be limited to geography. Second, understanding social context as space is the crucial step for bringing together spatial econometrics and social influence, since spatial weights matrices can be used to representing the proximities among individuals in a given space. Such spatial weights matrices can then be used in tests for spatial autocorrelation and in spatial regression models to assess whether there is evidence of a neighborhood effect, or social influence.

The relationship between spatial regression models and conventional models

In chapters 2 and 3, I compared and contrasted spatial regression models with analysis of variance, linear regression, contextual models, and social network models for studying social influence. Chapter 2 showed that the contextual model and social network model can be viewed as special cases of the spatial lag model and spatial-x model. Both the spatial-x model and social network model assume that the influence of social networks is exogenous; however, the spatial-x model allows all data points to be considered during estimation, while the social network model is based on a truncated sample.

Chapter 4 addressed the broader issue of using linear regression or ANOVA (as a special case of linear regression) versus spatial regression models. I demonstrated that while ANOVA can be considered a special case of the first-order autoregressive (FAR) model, neither simple linear regression nor multivariate linear regression can be considered spatial regression models. These results show that using a host of geographical, demographic, ideological, or other variables as control variables in a linear regression model is no substitute for using these same variables in a spatial weights matrix in a spatial regression model. Furthermore, I showed that while simple and multivariate linear regression models can be forced into special cases of spatial regression models by discretizing the continuous independent variables, the resulting discretized regression model is problematic for a variety of reasons.

In contrast to "garbage in, garbage out" models, spatial regression models can meaningfully account for geographical, demographic, ideological characteristics as part of an individual's social context. Instead of looking at the marginal effect of a single characteristic, researchers can define meaningful social contexts (*i.e.*, neighborhoods), use spatial regression models to represent the influence process, and estimate the impact on individual political behaviors and attitudes.

Using spatial econometrics to estimate social influence in individual attitudes toward political issues

Chapters 5, 6, and 7 provided empirical applications of spatial econometrics to the study of social influence on political attitudes among geographically, demographically, and ideologically proximate individuals.

Chapter 5 showed that there is evidence of social influence among geographical neighbors. Individuals who reside in the same congressional district mutually influence each other's attitudes toward two political issues: access to abortion and the government's role in securing jobs and a good standard of living.

Chapter 5 also illustrated the differences between spatial regression and dummy variable regression that were discussed in chapter 3. Using congressional districts as disjoint neighborhoods, I compared the results of estimating a spatial regression model and contextual model and demonstrated that the latter model constrains the neighborhood effect such that $\rho = 1$, while the former allows for its estimation.

Chapter 6 examined whether socially similar individuals mutually influence each other's political attitudes. Using Gower's dissimilarity coefficient as a measure of social proximity, I found that there is evidence of social influence for attitudes toward the issue of the government's role in securing jobs and a good standard of living and the issue of government spending and services.

Chapter 6 also illustrated the use of a discretized linear regression model based on demographically-defined disjoint neighborhoods. A comparison of the estimates of this model and the spatial regression model clarified the differences between the two models for studying social influence.

Chapter 7 is an analysis of mutual influence among ideologically proximate individuals. Using survey respondents' ideological placements of political candidates and political parties, I constructed a basic ideological space in which all respondents are located and used the distances within this space to determine the ideological proximity among the individuals. These distances were then used to construct spatial weights matrices representing the connectivities among the individuals in a basic ideological space. The results show that the ideologically-similar individuals mutually influence each other's attitudes toward the issues of defense spending, the tradeoff between environmental protection and jobs, and access to abortion.

CONTRIBUTIONS AND IMPLICATIONS

This dissertation is about whether spatial econometrics can be used to study social influence and it showed that it can be done. By equating social influence with neighborhood effect, by viewing social contexts as spaces, and by going beyond the limitations of geographical space, this dissertation showed that this task is possible. By highlighting the explanatory advantages of spatial regression and by clarifying the relationship between spatial regression models and linear regression models and its variants, this dissertation showed that using spatial econometrics to study social influence is a desirable one.

While the evidence of social influence uncovered in the empirical chapters is hardly overwhelming, the primary purpose of the chapters is to illustrate the use of spatial econometrics using survey data for understanding social influence in politics.

The spatial econometric approach used in this dissertation is an alternative to conventional methods of studying social influence, such as contextual analysis and ANOVA. It offers estimates of social influence but does not require special social network data, does not assume that social influence is a simple average of network members, and does not require respondents to be conscious of any influence. Furthermore, by leveraging the idea of Blau space and ideological space, it does not restrict social influence to geographical space. This study argued for and provided evidence of spatial dependence in political attitudes, thereby showing the viability of spatial econometrics for studying social influence.

DIRECTIONS FOR FUTURE RESEARCH

The results of this dissertation point to several possibilities for further research

First, the empirical applications in this dissertation were limited to cross-sectional data from the 2004 ANES. The next step is to extend the spatial regression analysis used in this dissertation by applying the spatial lag model to other ANES datasets. By comparing neighborhood effects for different years, we can see if these effects are consistent over time and whether changes in these effects are different for different issues.

Second, this dissertation considered only three types of spaces: geographical, demographic, and ideological. Since social space can be broadly defined and mutual influence is certainly not limited to geography, demography, or ideology, it may be informative to study mutual influence in spaces defined by other characteristics, such language similarity, online activity, or media markets.

Finally, the empirical applications in this dissertation assumed that there is no spatial heterogeneity, which refers to the lack of structural stability of political attitudes over space by way of different functional forms or systematically varying parameters (see Anselin, 1988). In substantive terms, this means assuming that the extent and size of social influence is the same across a space. This study looked only at spatial dependence and assumed that the neighborhood effects in geographical, Blau, and ideological spaces are uniform across those spaces; it is certainly plausible, however, to suspect that social influence might operate differently in different parts of a space. For example, it may be the case that more highly educated individuals are less mutually influential compared with less educated individuals.

Investigating spatial heterogeneity in political attitudes in different social contexts would provide a better understanding of the patterns of social influence.

APPENDIX: DESCRIPTIVE STATISTICS FOR POLITICAL ATTITUDES

Table A.1. Descriptive Statistics and Other Details for 8 Political Issues (7-point scale)

ISSUE	MEAN	ST. DEV.	VALID CASES	MISS-ING	MINIMUM	MAXIMUM
Government versus private medical insurance	3.66	1.92	1112	100	Government insurance plan	Private insurance plan
Environment versus jobs tradeoff	3.59	1.58	1019	193	Protect the environment, even if it costs jobs	Maintain/increase jobs and standard of living
Government's role in securing jobs and a good standard of living	4.21	1.78	1103	109	The government should see to jobs and standard of living	The government should let each person get ahead on his own
Government spending and services	4.52	1.59	1060	152	The government should provide fewer services	The government should provide many more services
Government assistance to blacks	4.54	1.79	1073	139	The government should help blacks	Blacks should help themselves
Role of women	1.92	1.47	1157	55	Women and men should have equal roles	Women's place is in the home
Defense spending	4.57	1.48	1061	151	Decrease defense spending	Increase defense spending
Access to abortion (4-point scale)	2.82	1.14	1055	157	Abortion should never be permitted	Abortion should always be permitted

REFERENCES

Achen, Christopher. 1992. "Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research." *Political Behavior* 14(3): 195-211.

Achen, Christopher. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5: 423-450.

Achen, Christopher. 2005. "Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong." *Conflict Management and Peace Science* 22(4): 327-339.

Anselin, Luc. 1995. "Local Indicators of Spatial Autocorrelation – LISA." *Geographical Analysis* 27(2): 93-115.

Anselin, Luc. 1998. *Spatial Econometrics: Methods and Models*. Norwell, MA: Kluwer Academic Publishers.

Anselin, Luc. 2006. "Spatial Econometrics." In T.C. Mills and K. Patterson, eds., *Palgrave Handbook of Econometrics: Volume 1, Econometrics Theory*. Basingstoke: Palgrave Macmillan, 901-941.

Ansolabehere, Stephen, Jonathan Rodden, and James Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference stability, Ideological Constraint, and Issue Voting." *American Politial Science Review* 102(2): 215-232.

Aristotle. 1995 [350 BC]. *Politics*. Trans. Ernest Barker. Oxford: Oxford University Press.

Asch, Solomon. 1951. "Effects of Group Pressure upon the Modification and Distortion of Judgment. In *Groups, Leadership, and Men*, ed. H. Guetzkow. Pittsburgh, PA: Carnegie Press.

Bargh, John and Tanya Chartrand. 1999. "The Unbearable Automaticity of Being." American Psychology 54(7): 462-479.

Beck, Nathaniel, Kristian Gleditsch, and Kyle Beardsley. 2006. "Space is More Than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50: 27-44.

Berelson, B.R., P.F. Lazarsfeld, and W.N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago, Illinois: University of Chicago Press.

Binder, Andrew, Kajsa Dalrymple, Dominique Brossard, and Dietram Scheufele. 2009. "The Soul of a Polarized Democracy: Testing Theoretical Linkages Between Talk and

Attitude Extremity During the 2004 Presidential Election." *Communication Research* 36: 315-340.

Blalock, Hubert. 1984. "Contextual-Effects Models: Theoretical and Methodological Issues." *Annual Review of Sociology* 10: 353-372.

Blau, Peter. 1957. "Formal Organization: Dimensions of Analysis." *American Journal of Sociology* 63(1): 58-69.

Boyd, Lawrence and Gudmund Iversen. 1979. *Contextual Analysis: Concepts and Statistical Techniques*. Belmont, CA: Wadsworth, Inc.

Brueckner, J.K. 2003. "Strategic Interaction Among Governments: An Overview of Empirical Studies." *International Regional Science Review* 26(2): 175-188.

Burt, Ronald. 1987. "Social Contagion and Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology* 92(6): 1287-1335.

Butler, David and Donald Stokes. 1974. *Political Change in Britain: The Evolution of Electoral Choice*. New York: St. Martin's.

Campbell, Angus, Warren Miller, and Philip Converse. 1960. *The American Voter*. New York: Wiley.

Chen, J. and J. Rodden. 2009. "Tobler's law, Urbanization, and Electoral Bias: Why Compact, Contiguous Districts are Bad for the Democrats." Working paper. Department of Political Science, Stanford University.

Cho, Wendy. 2003. "Contagion Effects and Ethnic Contribution Networks." *American Journal of Political Science* 47(2): 368-387.

Converse, Philip. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Its Discontents*, Edited by David Apter. New York: The Free Press of Glencoe.

Coughlin, C.C., Garrett, T.A., Hernández-Murillo, R. 2007. "Spatial Dependence in Models of State Fiscal Policy Convergence." ***Public Finance Review*** 35(3):361-84.

Cozier, Tamy. 2011. "'GOP is the New Black' claim Texas billboards." *The Grio*. Accessed November 30, 2011. http://www.thegrio.com/politics/billboard-the-gop-is-the-new-black.php.

Crano, William and Radmila Prislin. 2006. "Attitudes and Persuasion." Annual Review of Psychology 57: 345-374.

Davis, James, Joe Spaeth, and Carolyn Huson. 1961. "A Technique for Analyzing the Effects of Group Composition." *American Sociological Review* 26(2): 215-225.

Delli-Carpini, Michael and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. New Haven: Yale University Press.

Dow, Malcolm, Michael Burton, and Douglas White. 1982. "Network Autocorrelation: A Simulation Study of a Foundational Problem in Regression and Survey Research." *Social Networks* 4: 169-200.

Dow, Michael Burton, Douglas White, and Karl Reitz. 1984. "Galton's Problem as Network Autocorrelation." *American Ethnologist* 11(4): 754-770.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: HarperCollins.

Druckman, James. 2004. "Political Preference Formation: Competition, Deliberation, and the (Ir)Relevance of Framing Effects." *American Political Science Review* 98(4): 671-686.

Druckman, James and Arthur Lupia. 2000. "Preference Formation." *Annual Review of Political Science* 3: 1-24.

Elkins, Zachary and Beth Simmons. 2005. "On Waves, Clusters, and Diffusion: A Conceptual Framework." *The Annals of the American Academy of Political and Social Science* 598: 33-51.

Erbring, Lutz and Alice Young. 1979. "Individuals and Social Structure: Contextual Effects as Endogenous Feedback." *Sociological Methods and Research* 7(4): 396-430.

Franzese, Robert and Jude Hays. 2007a. "Empirical Models of Spatial Interdependence." Manuscript.

Franzese, Robert and Jude Hays. 2007b. "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15: 140-164.

Gleditsch, Kristian. 2002. A*ll International Politics is Local: The Diffusion of Conflict, Integration, and Democratization*. Ann Arbor: University of Michigan Press.

Gleditsch, Kristian and Michael Ward. 2007. "Diffusion and the Spread of Democratic Institutions," In F. Dobbins, G. Garret, B. Simmons, eds. *The Global Diffusion of Democracy and Markets*. Cambridge: Cambridge University Press.

Gleditsch, Kristian and Michael Ward. 2000. "War and Peace in Space and Time: The Role of Democratization," *International Studies Quarterly* 44(1): 1–29.

Gower, J.C. 1971. "A General Coefficient of Similarity and Some of Its Properties." *Biometrics* 27: 623-637.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6): 1420-1443.

Grossback, Lawrence, Sean Nicholson-Crotty, and David Peterson. 2004. "Ideology and Learning in Policy Diffusion." *American Politics Research* 32(5): 521-545.

Hinich, Melvin and Michael Munger. 1994. *Ideology and the Theory of Political Choice*. Ann Arbor: University of Michigan Press.

Hoaglin, David and Roy Welsch. 1978. "The Hat Matrix in Regression and ANOVA." *The American Statistician* 32(1): 17-22.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39(153): 41–57.

Huckfeldt, Robert. 1979. "Political Participation and the Neighborhood Social Context."

Huckfeldt, Robert. 1983a. "Social Contexts, Social Networks, and Urban Neighborhoods: Environmental Constraints Upon Friendship Choice." *American Journal of Sociology* 89: 651-669.

Huckfeldt, Robert. 1983b. "The Social Context of Political Change, Durability, Volatility, and Social Influence."

Huckfeldt, Robert. 1986. *Politics in Context: Assimilation and Conflict in Urban Neighborhoods*. New York: Agathon Press.

Huckfeldt, Robert. 2001. "The Social Communication of Political Expertise." *American Journal of Political Science* 45: 425-438.

Huckfeldt, Robert. 2007a. "Information, Persuasion, and Political Communication Networks." In *Oxford Handbook of Political Behavior,* R. Dalton and H.-D. Klingemann eds. Oxford, UK.

Huckfeldt, Robert. 2007b. "Unanimity, Discord, and the Communication of Public Opinion." *American Journal of Political Science* 51: 978-995.

Huckfeldt, Robert. 2009. "Interdependence, Density Dependence, and Networks in Politics." *American Politics Research* 37(5): 921-950.

Huckfeldt, Robert and Carol Kohfeld. 1989. *Race and the Decline of Class in American Politics*. Urbana: University of Illinois Press.

Huckfeldt, Robert and Jeanette Morehouse Mendez. 2008. "Moths, Flames, and Political Engagement: Managing Disagreement within Communication Networks." *Journal of Politics* 70(1): 83-96.

Huckfeldt, Robert and John Sprague. 1991. "Discussant Effects on Vote Choice: Intimacy, Structure, and Interdependence." *Journal of Politics* 53: 122-158.

Huckfeldt, Robert and John Sprague. 1995. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*. New York: Cambridge University Press.

Iverson, Gudmund. 1991. *Contextual Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Beverly Hills, CA: Sage Publications, Inc.

Jessee, Stephen. 2010. "Partisan Bias, Political Information and Spatial Voting in the 2008 Presidential Election." *The Journal of Politics* 72(2): 327-340.

Kaufman, L. and P.J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.

Kelejian, Harry, and Ingmar Prucha. 1998. "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance and Economics* 17(1): 99-121.

Kenny, Christopher. 1992. "Political Participation and Effects from the Social Environment." *American Journal of Political Science* 36(1): 259-267.

Kenny, Christopher. 1994. "The Microenvironment of Attitude Change." *The Journal of Politics* 56(3): 715-728.

Kinder, Donald. 1998. "Opinion and Action in the Realm of Politics." In *Handbook of Social Psychology*. Oxford: Oxford University Press.

King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton: Princeton University Press.

Klofstad, Casey. 2007. "Talk Leads to Recruitment: How Discussions about Politics and Current Events Increase Civic Participation." *Political Research Quarterly* 60(2): 180-191.

Lake, Ronald and Robert Huckfeldt. 1998. "Social Networks, Social Capital, and Political Participation." *Political Psychology* 19: 567-584.

Langton, Kenneth and Ronald Rapoport. 1975. "Social Structure, Social Context, and Partisan Mobilization: Urban Workers in Chile." *Comparative Political Studies* 8: 318-344.

Lazarsfeld, Paul, Bernard Berelson, and Hazel Gaudet. 1968. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press.

Lazer, David, Brian Rubineau, Carol Chetkovich, Nancy Katz, and Michael Neblo. 2010. "The Coevolution of Networks and Political Attitudes." *Political Communication* 27: 248-274.

LeSage, James. 1998. "Spatial Econometrics." Manuscript. University of Toledo. URL: http://www.spatial-econometrics.com/html/wbook.pdf.

Lipset, Seymour, Martin Trow, and James Coleman. 1956. *Union Democracy: The Internal Politics of the International Typographical Union*. New York: Free Press.

Lin, Tse-Min, Chin-En Wu, and Feng Yu Lee. 2006. "'Neighborhood' Influence on the Formation of National Identity in Taiwan: Spatial Regression with Disjoint Neighborhoods." *Political Research Quarterly* 59(1): 35-46.

Lorrain, F. and H.C. White. 1971. "Structural Equivalence of individuals in social networks." *Journal of Mathematical Sociology* 1: 49-80.

Lupia, Arthur. 2002. "Deliberation Disconnected: What It Takes to Improve Civic Competence." *Law and Contemporary Problems* 65(3): 133-150.

Luskin, Robert. 1987. "Measuring Political Sophistication." *American Journal of Political Science*. 31(4): 856-899.

Luskin, Robert, James Fishkin, and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in the U.K." *British Journal of Political Science* 32: 455-487.

Marsden, Peter. 1988. "Homogeneity in Confiding Relations." *Social Networks* 10: 57-76.

Marsden, Peter and Jeanne Hurlbert. 1987. "Small Networks and Selectivity Bias in the Analysis of Survey Network Data." *Social Networks* 9: 333-349.

McClurg, Scott. 2003. "Social Networks and Political Participation: The Role of Social Interaction in Explaining Political Participation." *Political Research Quarterly* 56(4): 449-464.

McClurg, Scott. 2006a. "The Electoral Relevance of Political Talk: Examining Disagreement and Expertise Effects in Social Networks on Political Participation." *American Journal of Political Science* 50(3): 737-754.

McClurg, Scott. 2006b. "Political Disagreement in Context: The Conditional Effect of Neighborhood Context, Disagreement and Political Talk on Electoral Participation." *Political Behavior* 28: 349-366.

McClurg, Scott. 2010a. "Methodological Challenges in the Study of Social Communication." Manuscript. Southern Illinois University Carbondale Department of Political Science Publications. URL: http://opensiuc.lib.siu.edu/ps_pubs/1.

McClurg, Scott. 2010b. "Porous Networks and Overlapping Contexts: Methodological Challenges in the Study of Social Communication and Political Behavior." In *Sourcebook for Political Communication Research: Methods, Measures, and Analytical Techniques*. Edited by Erik Bucy and R. Lance Holbert. New York: Routledge.

McPherson, Miller. 1983. "An Ecology of Affiliation." *American Sociological Review* 48(4): 519-532.

McPherson, Miller. 2004. "A Blau Space Primer: Prolegomenon to an Ecology of Affiliation." *Industrial and Corporate Change* 13(1): 263-280.

McPherson, Miller, Lynn Smith-Lovin, and Matthew Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks Over Two Decades." *American Sociological Review* 71: 353-375.

McPherson, Miller, Lynn Smith-Lovin, and James Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415-444.

Miller, Warren. 1956. "One-Party Politics and the Voter." *American Political Science Review* 50: 707-725.

Moody, James and Pamela Paxton. 2009. "Building Bridges: Linking Social Capital and Social Networks." *American Behavioral Scientist* 52☹11): 1491-1506.

Mutz, Diana. 1992. "Impersonal Influence: Effects of Representations of Public Opinion on Political Attitudes." *Political Behavior* 14(2): 89-122.

Mutz, Diana. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46(2): 838-855.

Mutz, Diana. 2002. "Cross-Cutting Social Networks: Testing Democratic Theory in Practice." *The American Political Science Review* 96(1): 111-126.

Nickerson, David. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102: 49-57.

North, Douglass. 1991. "Institutions." The Journal of Economic Perspectives 5(1): 97-112.

Ordeshook, Peter C. 1976. "The Spatial Theory of Elections: A Review and a Critique." In *Party Identification and Beyond*, Ian Budge, Ivor Crewe, and Dennis Farlie, eds. New York: Wiley.

Pattie, Charles and Ron Johnston. 1999. "Context, Conversation, and Conviction: Social Networks and Voting at the 1992 British General Election." *Political Studies* 47:877-889.

Pattie, Charles and Ron Johnston. 2000. "People Who Talk Together Vote Together: An Exploration of Contextual Effects in Great Britain." *Annals of the Association of American Geographers* 90(1): 41-66.

Poole, Keith. 1998. "Recovering a Basic Space From a Set of Issue Scales." *American Journal of Political Science* 42(3): 954:993.

Poole, Keith. 2011. "Recovering a Basic Space in R."

Putnam, Robert. 1966. "Political Attitudes and the Local Community." *The American Political Science Review* 60(3): 640-654.

Putnam. Robert. 1966. "Political Attitudes and the Local Community." *The American Political Science Review* 60(3): 640-654.

Rao, C. Radhakrishna, Helge Toutenburg Shalabh, and Christian Heumann. 2008. *Linear Models and Generalizations: Least Squares and Alternatives*. Springer Series in Statistics.

Rodden, Jonathan. 2009. "Federalism and Inter-Regional Redistribution." Working Paper. Barcelona Inst. Econ.

Rodden, Jonathan. 2010. "The Geographic Distribution of Political Preferences." *Annual Review of Political Science* 13: 321-340.

Sailer, Lee. 1978. "Structural Equivalence: Meaning and Definition, Computation and Application." *Social Networks* 1: 73-90.

Schelling, Thomas. 1978. *Micromotives and Macrobehavior*. New York: Norton.

Segal, David and Marshall Meyer. 1974. "The Social Context of Political Partisanship." In *Social Ecology*, ed. Mattei Dogan and Stein Rokkan. Cambridge: MIT Press, 217-32.

Sigelman, Lee, and Susan Welch. 1993. "The Contact Hypothesis Revisited: Black-White Interaction and Positive Racial Attitudes." *Social Forces* 71 (3): 781-95.

Simmons, Beth and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98 (1): 171-89.

Stipak, Brian and Carl Hensler. 1982. "Statistical Inference in Contextual Analysis." *American Journal of Political Science* 26(1): 151-175.

Sugiyama, Natasha. 2008. "Ideology and Networks: The Politics of Social Policy Diffusion in Brazil." *Latin Americans Research Review* 43(3): 82-108.

Sunstein, Cass. 2000. "Deliberative Trouble? Why Groups Go to Extremes." *The Yale Law Journal* 110(1): 71-119.

Taylor, Paul and Scott Keeter. 2010. "Millennials: A Portrait of Generation Next." Pew Research Center. URL: http://pewsocialtrends.org/files/2010/10/millennials-confident-connected-open-to-change.pdf.

Tingsten, Herbert. 1963. *Political Behavior: Studies in Election Statistics*. Totowa, New Jersey: Bedminster Press.

Ward, Michael and Kristian Gleditsch. 2002. "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace," *Political Analysis* 10(3): 244–60.

Ward, Michael and Kristian Gleditsch. 2008. *Spatial Regression Models*. London: Sage.

Werck, K., B. Heyndels, and B. Geys. 2006. "It Depends on Who You're Looking at: Neighbourhood Effects in Local Government Cultural Expenditures." 14th ACEI Conference on Cultural Economics (Vienna, July 2006).

Wiggins, Mark. 2011. "'GOP is the new black' claim East Austin billboards." KVUE ABC. Accessed November 30, 2011. http://www.kvue.com/home/GOP-is-the-new-black-claim-East-Austin-billboards-132744378.html.

Woods, Neal D. 2006. "Interstate Competition and Environmental Regulation: A Test of the Race-to-the-Bottom Thesis." *Social Science Quarterly* 87 (1): 174-89.

Volden, Craig, Michael Ting, and Daniel Carpenter. 2008. "A Formal Model of Learning and Policy Diffusion." Manuscript. URL: http://www.columbia.edu/~mmt2033/federalism.pdf.